

University of San Diego

Digital USD

Dissertations

Theses and Dissertations

2005-01-01

The Effect of Embedded Metacognitive Prompts and Probes on Students' Awareness in a Multimedia Lesson for Elementary School Students

Wendy Janine Parcel EdD
University of San Diego

Follow this and additional works at: <https://digital.sandiego.edu/dissertations>



Part of the [Leadership Studies Commons](#)

Digital USD Citation

Parcel, Wendy Janine EdD, "The Effect of Embedded Metacognitive Prompts and Probes on Students' Awareness in a Multimedia Lesson for Elementary School Students" (2005). *Dissertations*. 876.
<https://digital.sandiego.edu/dissertations/876>

This Dissertation: Open Access is brought to you for free and open access by the Theses and Dissertations at Digital USD. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital USD. For more information, please contact digital@san Diego.edu.

**THE EFFECT OF EMBEDDED METACOGNITIVE PROMPTS AND PROBES
ON STUDENTS' AWARENESS IN A MULTIMEDIA LESSON
FOR ELEMENTARY SCHOOL STUDENTS**

by

Wendy Janine Parcel

A dissertation submitted to the faculty of
San Diego State University and the University of San Diego
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

Dissertation Committee:

Brock S. Allen, Ph.D., San Diego State University
Jerome J. Ammer, Ph.D., University of San Diego
Marcie Bober, Ph.D., San Diego State University
John Watson, Ph.D., EyeCues

January 2005

Copyright © 2005

by

Wendy Janine Parcel

All Rights Reserved

DEDICATION

To my husband, Richard Hurshol Parcel, for always believing I can do anything I set my mind to. I've truly been blessed.

To my son, Maxwell Hurshol Parcel, for loving me and for believing that everyone's mommy goes to school. You and Daddy are "my guys."

To my parents, Dannie and Diana Carroll, for insisting that an education is important, and for reading to me as a child every evening before I went to sleep. Even without knowing it, you asked all the right metacognitive questions during the reading.

And to my sister, Dr. Tammy Lynne Prida. You are an inspiration and show that hard work and tenacity do make the difference.

ABSTRACT OF THE DISSERTATION

The Effect of Embedded Metacognitive Prompts and Probes on Students' Awareness in a
Multimedia Lesson for Elementary School Students

by

Wendy Janine Parcel

Doctor of Education

San Diego State University – University of San Diego, 2005

In a study he called *The Effect of Embedded Metacognitive Cues and Probes on Use of Learner Control Features in an On-line Lesson for Elementary Students*, Watson (2001) found that minimal prompting by an online tutorial increased 5th grade students' comprehension of how much they understood. While Watson's findings demonstrated a significant difference in the ability of prompted and non-prompted students to accurately predict their own performance on posttests, actual scores were not greater than those of control students until the results were analyzed by gender. The purpose of this study was to replicate and extend Watson's study to determine if the gender differences illuminated in the original study were replicable.

The extension was of two parts. The first called for administration of the Inventory of Metacognitive Self-Regulation (IMSR). The IMSR determines a student's unassisted level of metacognitive ability—his or her metacognitive trait level. While Watson's data collection looked only at students' momentary awareness of their own metacognition, the IMSR is a more refined measure of metacognition and allowed for additional analysis.

The second called for examination of how students interpret the metacognitive prompt pages. Watson found some evidence of gender differences in performance on the posttest, very possibly because boys and girls interpreted the prompts differently. The researcher added an experimental group and subjects were given a small incentive to do well on the posttest. The hypothesis was that the incentive would entice the boys to stay more focused on the task of scoring well on the posttest, instead of exploring the tutorial's user-control features.

The researcher administered the online tutorial to 147 fifth grade students at one of three different elementary schools. A multiple linear regression showed that all students predicted moderately well, with those in the prompted group not predicting any more accurately than students in the control group. Other statistics calculated yielded non-significant results. Post-hoc analysis showed students scored significantly different on the two posttest measures, although this was not true in the original study. While Watson's original findings were not supported, the questions he raised about whether metacognitive prompts increase metacognitive awareness and possible gender differences in prompt interpretation are valid questions worth pursuing in future research.

TABLE OF CONTENTS

	PAGE
ABSTRACT	v
LIST OF TABLES	ix
LIST OF FIGURES	xi
ACKNOWLEDGEMENTS	xii
CHAPTER	
1 INTRODUCTION.....	1
Statement of the Problem and Purpose of the Study	3
Research Questions and Hypotheses	5
Main Research Question	5
Research Question Two	5
Research Question Three	6
Research Question Four	7
Research Question Five	8
Significance and Contribution of Planned Research	8
Definition of Terms	10
2 REVIEW OF THE LITERATURE.....	11
History of Metacognition.....	11
Metacognition in Children	12
Metacognitive Beginnings in Children.....	13
Metacognition, Children, and Academic Achievement.....	14
Assessment Tool Literature	18
Inventory of Metacognitive Self-Regulation.....	18
Motivation Scale: Student Opinion Scale	19
Summary	20
3 METHODOLOGY	21
Review of the Research Questions	21
Context of the Study and Description of Student Populations.....	21

Population Sample.....	26
Instrumentation.....	27
Online Tutorial	27
Tutorial Posttests	28
Inventory of Metacognitive Self Regulation.....	28
Subject Recruitment and Random Assignment.....	29
4 RESULTS	30
Subjects.....	30
Posttest Instrument Reliability and Validity	31
Data Collection.....	31
Research Question One.....	33
Research Questions Two and Three	35
Research Question Four	36
Quantitative Analysis.....	38
Qualitative Analysis	39
Research Question Five	41
IMSR.....	43
Additional Data Analysis to Compare Posttests.....	44
Difference Between the Two Posttests.....	44
Analysis for Posttest One Only.....	45
5 DISCUSSION	47
Main Research Questions	47
Research Questions Two and Three	48
Research Question Four.....	49
Research Question Five	50
Strengths of the Study.....	51
Limitations of the Study.....	52
Tutorial Concerns	52
Posttests Concerns	53
Subjects' Metacognitive Abilities.....	54
Future Research.....	54
Tutorial Redesign and Subject Selection	55

Isolated Redesign.....	55
Metacognitive Level and Appropriate Interventions	55

APPENDICES

A PARENT PERMISSION SLIP	61
B STUDENT PERMISSION SLIP.....	65
C IMSR	68
D PERMISSION TO USE THE IMSR.....	71
E STUDENT OPINION SCALE	73
F INTERVIEW QUESTIONS	75
G INSTRUCTION FOR SUBJECTS	77
H STUDENT LOGIN LETTERS.....	79

LIST OF TABLES

	PAGE
Table 1. Subject Groupings	5
Table 2. Rank Order of the Mean Scores	7
Table 3. IMSR Constructs	19
Table 4. Question Summary	22
Table 5. Casillas Student Ethnicity	24
Table 6. Hilltop Drive Student Ethnicity	25
Table 7. Parkview Student Ethnicity	26
Table 8. School Comparisons.....	26
Table 9: Model Summary for Posttest One.....	34
Table 10: Regression Analysis Summary for Variables Predicting the Interaction Effect Between Students' Score Prediction on Posttest One and Prompted Status (N = 146)	34
Table 11: Model Summary for Posttest Two	35
Table 12: Regression Analysis Summary for Variables Predicting the Interaction Effect Between Students' Prediction of Scores on Posttest Two and Prompted Status (N = 146)	35
Table 13. Descriptive Statistics for Question Three	37
Table 14. Test of Between-Subjects Effects for Question Three	38
Table 15. Group Statistics	39
Table 16. Independent samples <i>t</i> test	39
Table 17. Tests of Between-Subjects Effects: Total Incentive.....	40
Table 18. Tests of Between-Subjects Effects: Importance Scale	41
Table 19. Tests of Between-Subjects Effects: Effort Scale	42
Table 20. Crosstabulation.....	42
Table 21. Correlation of IMSR Constructs Three and Four to Posttest Scores.....	43
Table 22. ANCOVA	44
Table 23. Paired Samples Statistics for Posttests One and Two	44
Table 24. Paired Samples Correlations for Posttests One and Two	45

Table 25. Paired Samples Test for Posttests One and Two.....	45
Table 26. Test of Between-Subjects Effects, Posttest One Only.....	46

LIST OF FIGURES

	PAGE
Figure 1. Metacognitive Prompt Screen.....	4
Figure 2. Guiding Question Tutorial Screen	53

ACKNOWLEDGEMENTS

It would not have been possible for me to complete this dissertation without the support and guidance of several people. I was blessed to have the perfect committee. Brock Allen, Jerry Ammer, Marcie Bober, and John Watson mentored me along the way and were always prompt, professional and a joy to work with.

I thank my chair, Dr. Brock Allen, for believing in me. I'm grateful that Brock called forth my best effort, and yet had patience with my writing and thinking. Seeing the happy look on Brock's face at the end of the data analysis will probably stay with me for life.

Thank you to Dr. Jerry Ammer, for helping me think about *all* students and for pushing me to keep expanding my knowledge base. Jerry listened patiently and helped me work out the story being told by the data.

Thank you to Dr. Marcie Bober who encouraged me to enter the doctoral program five years ago. Marcie is always there for me when I need someone to talk to, and has helped me become a better statistician than I ever thought I'd be.

Thank you to Dr. John Watson, without whom I would not have had a dissertation to replicate and who made himself available to me at all times and answered my many questions.

Thank you to Rusty Meike, for presenting me with the "certificate of confidence" several years ago when I thought I was about to break, and for reading and editing, and re-reading and re-editing, the dissertation. Thank you for being my friend.

Marc Pastor, is truly the nicest guy I've ever met. Over lunch one day Marc volunteered to recreate the tutorial for me and would not accept any form of payment—a huge worry of mine was suddenly gone. Thank you Marc.

While the dissertation has been a challenge, it does not even begin to compare with becoming a mother. Thank you to all the moms on the ABD_Moms Listserv for providing encouragement, support, and for sharing your unique, but common stories.

And finally, thank you to my hubby, Richard.

CHAPTER 1

INTRODUCTION

Metacognition is thinking about thinking. While the application of what is known about metacognition has been useful in areas such as education and eyewitness testimony, there is not one dominant theory or framework of metacognition upon which all metacognitive researchers agree. Yet, broad definitions such as Flavell's, "metacognition refers to one's knowledge concerning one's own cognitive processes and products or anything related to them," (1976, p. 232) are generally the starting point of metacognitive research.

Adult learners who are more metacognitively aware score higher on standardized tests than those who are less metacognitively aware (Schoenfield, 1987). Not only are those with greater metacognitive abilities able to identify what problems need to be solved, but also they are able to monitor their responses and their proposed solutions to determine that they are meeting their goals. Some researchers have even suggested that high metacognitive ability can compensate for lack of content knowledge when taking standardized tests (Howard, McGee, Shia & Hong; 2001b; Swanson, 1990).

Flavell, as early as 1979, argued that metacognition helps explain why children of varying ages deal with learning tasks in different ways. Sperling, Walls and Hill (2000) in fact showed that children as young as two or three demonstrate some metacognitive ability, and many studies (Kreutzer, Leonard & Flavell, 1975; Kuhn, 2000; Schneider & Lockl, 2002; Schneider & Pressley, 1997) suggest that these abilities tend to improve with age. In the last thirty years, researchers have conducted studies concerning children's metacognitive awareness as they deal with reading (and to some extent, math), but the research on the use of computer environments to enhance children's metacognition in any area is inconsistent and incomplete. There is not much of a theoretical base for prior research, and until recently, there was not even a widely validated inventory to measure children's metacognition.

According to Brown (1978), there are two components to metacognition, Knowledge of Cognition and Regulation of Cognition. Knowledge of Cognition refers to how much a

learner understands about his own memory organization and the way he learns, while Regulation of Cognition is defined as how well learners can regulate their own memory learning. Brown believes that metacognitive effects on cognitive regulation are more important than other metacognitive functions, and that “children do not monitor well and often fail to make appropriate executive decisions” (Schneider & Pressley, 1997, p. 194). Kuhn suggests that children are not alone in their frequent failure to regulate cognition. “We need to know more about how it (metacognition) develops and how it comes to regulate first-order cognition, or, very often, fails to do so. The fact that such failure is a common occurrence raises what is perhaps the most consequential question in need of more investigation: How can metacognitive development be facilitated?” (p. 180).

Seasoned educators and parents know that one way to facilitate metacognitive development is to scaffold the learning environment for the student. Vygotsky (1978) called the difference between what a child can do with help and what he or she can do without guidance the zone of proximal development. When students are in this zone of proximal development, teachers will ask a question and then pause to give the student time to think. If, at the end of the pause, the student does not know the answer, a good teacher does not provide the solution, but instead provides clues to help facilitate the thinking process. In this environment, the clue might be as simple—reminding the student about prior knowledge. Or, the clue might be as complicated as providing a similar question and then showing the strategy used to solve the other problem in hopes that the student can then transfer the strategy to the original question. What the teacher is doing for the student is acting as an external regulator of cognition. Often, the student has the cognitive knowledge to solve the problem, but his cognitive regulator is not selecting the appropriate strategy activating the correct prior knowledge. With guidance and practice, a student learns to regulate his own cognition and is then able not only to solve the problem on his own, but also to determine that there is a problem to be solved and to evaluate whether his solution is likely to be correct.

Typically, elementary classrooms contain students of varying ability levels. It is seldom possible for an educator to facilitate metacognitive behavior for each individual student on a regular basis, regardless of the ways he or she facilitates instruction. One way to provide help and scaffolding for students is with computer-based instructional software.

Ninety percent (48 million) of children between the ages of 5 and 17 use computers (U.S. Department of Commerce). As computers become ubiquitous in classrooms, and as software designers understand that computer programs can not only teach curriculum content, but also prompt and scaffold metacognitive behavior, it is important to know the effects metacognitive prompting have on students when included in educational software.

STATEMENT OF THE PROBLEM AND PURPOSE OF THE STUDY

Little is known about the effects of metacognitive prompts in K-6 educational software. This dearth of research motivated the current study. The researcher located only one well-conducted and documented experiment that directly manipulated the metacognitive experiences of students and then subsequently observed their cognitive performance. In *The Effect of Embedded Metacognitive Cues and Probes on Use of Learner Control Features in an On-line Lesson for Elementary Students*, conducted in 2001, Watson found that even minimal amounts of prompting by the tutorial increased students' perception of how well they would perform.

The current study replicated and extended Watson's study by examining the effect of metacognitive cues on learner interactions as they complete an online tutorial about weather and cloud formation.

Watson used two groups for his study. The experimental group received metacognitive prompts while the control group did not. Both groups read and clicked through an online tutorial that introduced students to basic concepts related to weather phenomena, such as types of clouds. Students completed two online posttests, one midway through the tutorial and another at the end. The screen preceding the first posttest presented a metacognitive probe—a question asking students to predict how well they would do on that posttest by selecting one of three choices. The ratio of students' predictions to their posttest scores was then used as a measure of their metacognitive awareness. Directly before the prediction screen, the prompted group saw a screen like in Figure 1.

Watson demonstrated a significant difference in the ability of the prompted and non-prompted students to accurately predict their own performance on the posttest measures. The prompted students demonstrated a significantly greater awareness of how much they understood, suggesting that the prompt and probe momentarily increased their metacognitive

awareness. While the aggregate scores were not significantly different, a post-hoc analysis revealed gender differences as follows:

- Girls in the experimental group scored higher on the posttest than girls in the control group.
- Boys in the experimental group scored lower on the posttest than boys in the control group.

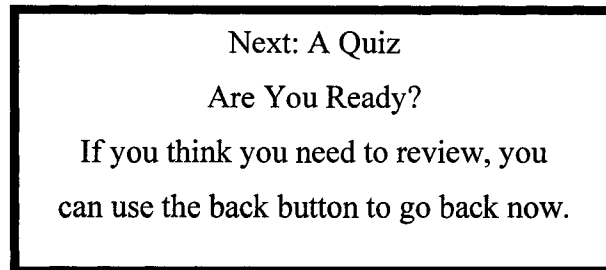


Figure 1. Metacognitive prompt screen.

In essence, the girls appeared to have benefited from the metacognitive prompts while the boys had reduced posttest boys performance.

Watson's findings—that metacognitive prompts increased metacognitive awareness and that metacognition was not correlated with student achievement—are consistent with other empirical findings (Sperling, Howard, Miller, & Murphy, 2002; Swanson, 1990). A replication and extension study was warranted to determine if the gender differences illuminated in the original study were replicable.

The first extension used the Inventory of Metacognitive Self-Regulation (IMSR) (Howard, McGee, Shia, & Hong, 2000a). Validated with hundreds of subjects in grades 5-12, it measures five facets of metacognitive self-regulation: knowledge of cognition, objectivity, problem representation, subtask monitoring, and evaluation. The IMSR determines a student's unassisted level of metacognitive ability—his or her metacognitive trait level. While Watson's data collection looked only at students' momentary awareness of their own metacognition, the IMSR is a more refined measure of metacognition and allowed for additional analysis.

The second extension examined how students interpreted the metacognitive prompt pages. Watson(2001) found some evidence of performance differences by gender on the posttest, very possibly because boys and girls interpreted the prompts differently. In order to see if this difference could be accounted for, an experimental group was added and the

subjects were given a small incentive (money) to do well on the posttest. The belief was that the incentive would entice the boys to stay more focused on the task of scoring well on the posttest, instead of exploring the user-control features of the tutorial.

Table 1. Subject Groupings

Group 1:	Control (No metacognitive cue/prompt. No incentive.)
Group 2:	Experimental (No metacognitive cue/prompt, but with incentive)
Group 3:	Experimental (With metacognitive cue/prompt. No incentive.)
Group 4:	Experimental (With metacognitive cue/prompt (as in group 1) <i>and</i> an incentive to do well on the posttest.)

RESEARCH QUESTIONS AND HYPOTHESES

Main Research Question

Do metacognitive prompts increase student awareness of their own knowledge? This question is based on findings from: (a) Watson's original (2001) study suggesting that even minimal prompting increased student understanding of how much they understood, and (b) a review of other studies suggesting that various interventions, such as scaffolding and prompting, can improve metacognitive performance. Therefore, the research hypothesis is that learners who receive metacognitive prompts will better predict their actual scores on posttest measures of comprehension than learners in the control group. The null hypothesis is that there will not be a significant difference between the experimental and control groups. The treatment in this study employed prompting, but not scaffolding.

Research Question Two

Do metacognitive prompts increase student performance on posttest measures of recall of tutorial content? Watson (2001) hypothesized that prompted students would be more likely to revisit and attend to portions of the tutorial than non-prompted students, and would therefore score higher on posttest measures. Although he found no overall difference between prompted and non-prompted students in use of navigational controls, path analysis revealed a significant difference in the use of the "back button" by prompted students (Watson, p. 55). However, Watson found no significant difference between the two groups on posttest scores of comprehension.

Sperling and colleagues (2002) also found little evidence to support assumptions that metacognition is positively related to achievement or that metacognition contributes to achievement. Although Sperling and her colleagues did not explain this lack of evidence, it suggests that if metacognitive abilities are to contribute to actual academic achievement, the student must have an incentive to employ those metacognitive abilities.

Other factors may have contributed to Watson's (2001) inability to find performance differences between prompted and non-prompted students. The metacognitive prompts employed in Watson's study were fairly low-level, reminding students to perform two of the sub-skills of metacognition: monitoring and evaluation. Such reminders may not change the behaviors of students who are already metacognitively aware. In addition, Watson's posttest comprehension measure, a multiple-choice test with only ten items and a split-half reliability of .67, may not have been sensitive enough to detect a difference between prompted and non-prompted students. Therefore, the current replication study employed: (a) more posttest items, and (b) takes additional steps to assure alignment of those items to the learning outcomes.

The research hypothesis predicts that experimental and control groups receiving treatments similar to those employed in the Watson study will yield similar results: no significant difference between prompted and non-prompted students.

Research Question Three

Is gender a factor in the effect of prompts on posttest performance? Watson's post-hoc exploratory data analysis suggests this question. Watson (2001) found that students in the experimental group scored the same on posttest measures of tutorial content as students in the control group—a surprising finding which led him to conduct a post-hoc data analysis by gender that revealed an interesting gender interaction, e.g., the girls in the experimental group scored higher than girls in the control group, while the boys in the experimental group scored lower than the boys in the control group. This interaction, however, was not significant after making adjustments for multiple exploratory tests in the post-hoc analysis.

The study investigated the possibility that girls and boys do indeed respond differently to the metacognitive prompts, but in the context of an underlying theoretical assumption that boys in Watson's study tended to disregard tutorial content while exploring navigation

functions because they were not incentive to achieve well on the posttest. The research hypothesis for Question Three is that posttest scores for boys who are promised rewards will not differ significantly from girls promised similar rewards, whereas scores for boys who are not promised rewards will be significantly lower than girls not promised rewards (See Table 2).

- When no external incentive is provided, the rank order of the mean scores (highest to lowest) will be as follows: prompted girls, non-prompted girls, non-prompted boys and prompted boys.
- When external incentive is provided, prompted students, both boys and girls, will score significantly higher on posttest measures than non-prompted students.

Table 2. Rank Order of the Mean Scores

	Girls Non-prompted	Boys Non-prompted	Girls Prompted	Boys Prompted
No Incentive	2	3	1	4
Incentive	2	2	1	1

*Gray row is the replication portion of the study

Research Question Four

Are there differences reported by boys and girls, both prompted and non-prompted, about their behaviors and thinking regarding review of tutorial content following announcement of the posttest? Watson did not address this question with interviews or other self-report data. He speculated that boys and girls interpreted metacognitive prompts in unique ways, and therefore responded to them in dissimilar fashions leading to the differences in posttest scores. In a study of computer software and gender, Caftori (1996) found that while 83% of girls indicated that they enjoyed mastering a computer game, only 22% of the boys so indicated (n.p.). Caftori noted that, “it seems as though girls in general, do not feel comfortable with new software, whereas it is not such a deterrent for boys” (n.p.). She raised the possibility that boys are used to being in strange new places, because more of them are allowed to explore new situations, and therefore do not see new software programs as “obstacles” the way girls do. Caftori also reported that many boys, but no girls, reported that they “don’t like to follow directions” when using computer software.

Caftori’s research is consistent with a study that McKean, Hoffman and Allen (2000) conducted. In that investigation, boys found and manipulated many more objects than girls

when using a virtual learning environment about missions. Since there appears to be a tendency for boys to explore a software program more than girls and not stay as focused on the task, the hypotheses are:

- Boys who receive an incentive to perform well on the posttest measures, will report staying focused on maximizing their posttest performance, while boys who do not receive an incentive will report greater exploration of the navigational controls.
- Girls will not report any difference in behavior, whether or not they receive an incentive.

Research Question Five

What is the relationship between prior metacognitive ability and the effectiveness of metacognitive prompts? Watson did not address this issue and did not identify instruments that might have been used to measure a child's metacognitive abilities. The literature (Allen & Merrill, 1985, 1987; Mevarech, 1999) suggests that metacognitive prompts are more likely to enhance performance of students who are unlikely to employ metacognitive skills on their own. The underlying assumption is that students with high levels of metacognitive ability or aptitude are already able to employ their metacognitive abilities to study materials and prepare to answer questions, and therefore are unlikely to benefit as much from prompting as are students with lower metacognitive abilities. As part of this extension study, students completed an inventory of metacognitive awareness and skills based on the work of Howard, McGee, Shia and Hong (2000a). Their inventory includes a number of sub-scales, including Knowledge of Cognition, Objectivity, Problem Representation, Subtask Monitoring and Evaluation. An exploratory analysis was conducted to determine whether and to what extent factors based on these subscales, considered as independent variables, might account for variance in the measures associated with Research Questions 1-4, considered as dependent measures.

SIGNIFICANCE AND CONTRIBUTION OF PLANNED RESEARCH

One goal of education is to promote and develop self-regulated learners. Kuhn asserted, "We lack sufficient research observing individuals engaged in the process of acquiring new knowledge" (2000, p. 180). While many teachers and parents are directly concerned with the education of children today, future employers of these children also care

about the quality of education they are receiving (Packer, 2002; SCANS 2000 Center, 2001; Washington Roundtable, 1998). While children's scores on high-stakes tests are important, the ability to think and reason well is also a desirable quality (Packer, 2002; SCANS 2000 Center, 2001; Washington Roundtable, 1998). There is a link between a person's metacognitive ability and his success on many school related tasks, including the abilities to problem-solve and read effectively (Delclos & Harrington, 1991; Jacobs & Paris, 1987; Palincsar & Brown, 1984; Swanson, 1990). In addition, enhancing a student's metacognitive abilities increases the likelihood that she will perform better on both standardized tests *and* real-world problems (Davidson, Deuser, & Sternberg, 1994; White & Frederiksen, 1998). The power of metacognition is being able to select the appropriate strategy when needed. Still, students' independent use of metacognitive strategies is gradual in development (Scruggs, Mastropieri, Monson, & Jorgenson, 1985). Hennessey (1999, n.p) wrote that, "even in the best constructivist learning environment metacognition does not simply happen, it must be explicitly promoted," and that "changes in metacognitive sophistication can be gained by actively engaging in the process". Mevarech (1999), too, asserts that students need to be "trained to self-monitor their learning as they solve problems". The reality is, however that most teachers lack the time to provide continuous, sustained metacognitive coaching.

Computers are one means for teachers and parents to provide individualized instruction for students. Yet, while reports such as *A Nation Online* (U.S. Department of Commerce, 2002) and the Kaiser Family Foundation's *Kids & Media* (Roberts, Foehr, Rideout & Brodie, 1999) show students today are spending a great deal of time on computers, especially in the form of Internet use and games, researcher have not explored the effects of simply having a few metacognitive prompts placed in an online tutorial or lesson. As young learners use computers more, understanding their goals and educational needs is becoming increasingly important. This study will lead to a greater understanding of how students can benefit from metacognitive prompts in an online tutorial, and which students will benefit the most from these prompts. Software developers, educators and parents will all profit from the result of this study.

Educators and parents who understand the role that metacognition plays both in school and throughout life will be interested in seeing whether the computer can help teach metacognition. Metacognition has important applications in the field of education because of

its central role in problem solving and learning (Flavell, Miller, & Miller, 2002). If metacognition can be scaffolded using a computer, teachers will have one more way to help meet students' individual metacognitive needs. Teachers and parents must learn how to choose appropriate software for an individual child's metacognitive training needs.

With an understanding how metacognition can be scaffolded in online learning environments, educational software developers will be able to build better systems that cater to the needs of children.

DEFINITION OF TERMS

The following is a list of terms used in the study.

- **Metacognition:** The process of thinking about thinking including monitoring and self-regulation of cognitive processes.
- **Metacognitive prompts:** A metacognitive cue is a device to signal a person to reflect on his or her thought process. In a classroom, this may be an action by the instructor that induces a learner to reflect on the process of learning. Within a computer-based training environment, the cue may be a visual or audio signal within the software.
- **Metacognitive probes:** A metacognitive probe is a device within computer-based training or an action by an instructor that attempts to elicit a response from the learner to determine their current metacognitive state.
- **Regulation of cognition:** The ability of learners to regulate their own memory learning.

CHAPTER TWO

REVIEW OF THE LITERATURE

“The term metacognition has been used to describe our knowledge about how we perceive, remember, think, and act--that is, what we know about what we know” (Metcalfe & Shimamura, 1994, p. xi). This definition is broad enough to be accepted by many researchers studying metacognition, and simple enough to be understood the educated public. Definitions for metacognition in the broadest sense, such as this one, are seldom refuted. Yet, since Flavell introduced the concept of metacognition in 1976, researchers have been struggling with a more comprehensive definition.

In their book, *Metacognition: Knowing About Knowing*, Metcalfe and Shimamura (1994) outline the general trend of discourse on theory. Theoretical models of the causes and functions of metacognition have been proposed and tested against one another and, “although this...has produced rich and varied sets of empirical findings, experimental paradigms, and modeling techniques, it has not produced dominant theories or frameworks that expand on their predecessors. This failure to produce theories and frameworks that encompass the findings of prior decades is undoubtedly an important factor for the relatively slow rate of cumulative progress” (Nelson, & Narens 1994, p.3).

Although there is no one dominant theory or framework that includes in a logical manner all that is known about metacognition, writers acknowledge two major frameworks, one by Flavell (Flavell, 1979; Flavell et. al., 2002) and one by Brown (Brown, 1978; Brown, & Brown, 1984), and most current metacognitive researchers adopt one these as a basis for additional study.

HISTORY OF METACOGNITION

In 1976, Flavell introduced the concept of metacognition in the context of developmental psychology and research on metamemory (Simons, 1996). He defined metacognition then as “one’s knowledge concerning one’s own cognitive processes and products or anything related to them” (Flavell, 1976, p. 232). Flavell thought of metacognition as encompassing metacognitive knowledge and metacognitive experiences. In

1979 he expanded his definition, declaring, “I believe that the monitoring of a wide variety of cognitive enterprises occurs through the actions of and interactions among four classes phenomena: (a) metacognitive knowledge, (b) metacognitive experiences, (c) goals (or tasks), and (d) actions (or strategies)” (p. 906). Finally, in the 1990’s Flavell reorganized and wrote, “metacognition refers to metacognitive knowledge (persons, tasks, and strategies) and to metacognitive monitoring” (Flavell, 1999, n.p.).

Shortly after Flavell first published his thoughts on metacognition, Ann Brown proposed a somewhat different model. In 1978, Brown suggested that there were two components of metacognition—knowledge of cognition and regulation of cognition (Brown, 1978). *Knowledge of cognition* included declarative, procedural, and conditional knowledge, while *regulation of cognition* included planning, monitoring and evaluation. Schneider and Pressley, in their book *Memory Development Between Two and Twenty*, suggest that “an impression that could be gleaned from some of Flavell’s early research was that a lot of metacognitive development was complete by age 8 or 9” (Schneider & Pressley, 1997, p. 193). They explain that one motivation for Brown’s reconceptualization of metacognition was to “counteract” this impression that they suggested was the result of studying “isolated pieces” of children’s metacognition (Schneider & Pressley, 1997). They add that Brown’s frame of reference “was the competent information processor, one possessing an efficient ‘executive’ that regulated cognitive processes” (Schneider & Pressley, 1997, p. 194). These functions, when fully developed: analyze new problems, select strategies, monitor success and failures to revamp for next time, and know when it does or does not know something. Brown’s research clearly made the case that metacognitive abilities develop slowly during school years, and never reach full capacity in some adults.

METACOGNITION IN CHILDREN

While both Flavell and Brown’s interest in metacognition included conducting research with children, most metacognitive studies have focused on adults—probably because adults are easier to research than children, both in respect to gaining access through human subjects and in terms of having the linguistic capabilities to explain themselves. Still, there is a growing body of research examining when metacognitive constructs appear in children and how metacognition effects thinking skills and school success in general.

In 1999 Flavell wrote *Cognitive Development: Children's Knowledge About the Mind* in which he explains that there have been three main “waves” of research on the development of children's knowledge about the mind. The first wave of research, he says, predates him and stems from the work of Piaget. Piaget and his wife showed that young children do not know that they have perspectives and so they don't understand that other people might have different perspectives. He demonstrated that even after children have acquired this knowledge of what is now called the “other mind,” they need still more time to acquire the skill of understanding their own perspective as different from other's perspectives.

In the 1970s, the second wave began with Flavell's work on metacognition and includes the work of many other researchers, including that of Ann Brown. Flavell, when speaking about this second wave says, “metacognition includes knowledge about the nature of people as cognizers, about the nature of different cognitive tasks, and about possible strategies that can be applied to the solution of different tasks. It also includes executive skills for monitoring and regulating one's cognitive activities” (Flavell, 1999, n.p.).

Finally, Flavell says that in the third wave, beginning in the 1980s and continuing to this day, theory of the mind and metamemory studies dominate the research. Clearly in studies with younger children, those six and under, this is the case. Kuhn has said that, “over the past decade, the wave of research on children's understanding of the mind has been valuable in highlighting the earliest forms of metacognition” and that “by age 3, children have acquired some awareness of themselves and others as knowers” (Kuhn, 2000, p. 178). As these metacognitive abilities increase and a person has greater control of regulation over their own metacognition, they become more effective at using their metacognitive abilities. Kuhn has also said that, “young children's dawning awareness of mental functions lies at one end of a developmental progression that eventuates in complex metaknowing capabilities that many adults do not answer” (Kuhn, 2000, p. 178).

METACOGNITIVE BEGINNINGS IN CHILDREN

Researchers have shown that preschoolers know some things about thinking such as: only people (not plants, rocks, etc.) engage in thinking, thoughts and images are internal, and that desires and thinking can take as their objects nonrepresent and even nonreal things.

Children are not, however, good at inferring what a person is thinking about, often not even realizing that someone must be thinking (Flavell, 1999).

It is also known that as children get older, their ability to predict how well they will do or how well they have done increases, although even for adults it is more difficult to predict how well you will do versus how well you have done. Brown showed that, “the ability to accurately assess performance after a response is made contrasts sharply with the ability to predict accuracy prior to a retrieval attempt. Predicting in advance of responding requires the ability to imagine cognitive acts that have not yet occurred. There is considerable evidence that such acts of imagination are more difficult for the young child” (Brown, 1978, p. 86). Other studies concur with these findings saying that there is a strong tendency toward improvement at the 4th and 5th grade level, with younger grade children experiencing, “wishful thinking” when asked to predict how well they will do (Schneider & Pressley, 1997).

METACOGNITION, CHILDREN, AND ACADEMIC ACHIEVEMENT

Kurtz and Borkowski (1985) studied metacognition as it relates to children’s reading abilities hoping to discover if metacognitive training would be academically beneficial to children’s abilities to summarize. Their study was comprised of 130 children in grades 4-6 who were divided into two experimental and one control group. The first experimental group received summarizing instruction, while the second experimental group was provided with metacognitive training in addition to the summarizing instruction. The control group received neither metacognitive nor summarizing instruction.

They found that children who received both summarization and metacognitive training had superior performance. They believed this was because metacognitive ability is a causal antecedent of later strategy acquisition. In other words, children who received metacognitive training were better able to understand summarizing strategies and know when to apply them. Kurtz and Borkowski said, “this highlights the dual importance of metacognitive knowledge as a precursor of later strategy acquisition and metacognitive skills as the ‘executor’ for lower-level strategies” (Kurtz & Borkowski, 1985, p. 1). They also say, “reading programs should not underestimate the importance of preexisting metacognitive deficiencies. Although these deficiencies are often qualitatively different from the academic

skills being instructed, they may be critical for producing long-term training successes” (Kurtz & Borkowski, 1987, p. 146).

In 1999, Mevarech also found that students who were given metacognitive training, in addition to strategy training, performed better than those given only strategy training. The focus of Mevarech’s study was on solving compare problems, which include relational terms such as more or less. She said that she selected compare problems because often the “language was inconsistent in them making the problems difficult to solve,” and also because they employ both “linguistic and nonlinguistic elements and thus may require various problem-solving skills” (Mevarech, 1999, n.p.).

As in the Kurtz and Borkowski study, Mevarech employed three groups. One was taught metacognitive and strategy application, one was taught only strategy application, and one received no training. The 174 students, 86 boys and 88 girls, were all seventh graders in Israel. A 20-item paper-and-pencil test administered after instruction was used to assess students’ ability to solve two-step compare problems. (Mevarech, 1999)

In her findings, Mevarech showed that students in the group that got metacognitive training did the best. She conducted a detailed analysis that indicated, “although both lower and higher achievers benefited from the metacognitive training, none of the lower achievers who were exposed only to the strategy instruction was able to solve the transfer problems” (Mevarech, 1999, n.p.). This means that the high achievers in the group that received only strategy training were able to solve some of the transfer problems, and probably employed previously acquired metacognitive strategies to assist them. It appears that while teaching metacognitive strategies is beneficial to high achievers, it is absolutely crucial for low achievers.

Sternberg examined the relationship of IQ to metacognition using fourth, fifth and sixth graders of either high or average intelligence as the study participants. Subjects received insight problems with and without cueing of one of the three kinds of insight. It was found that the highly intelligent children “spontaneously produced the three types of insights required to solve the problems; therefore their performance improved very little when each type of insight was cued” (Davidson, Deuser, & Sternberg, 1994, p. 222). The children who benefited most from the cueing were those of average intelligence. This would seem to suggest that not only is metacognitive ability linked to IQ and useful when solving problems,

but that in addition to directly teaching metacognitive strategies one can also provide metacognitive scaffolding in the form of prompts or cues to help children, and presumably adults. By scaffolding metacognition often enough, there is a great likelihood that as with other types of scaffolded learning, the student will eventually no longer need the scaffolding. Kuhn says that, “a good new strategy gradually appears more frequently if a child has the metacognitive self-regulatory skills to detect whether the strategy is useful for the task at hand” (2000, p. 180).

Researchers in the three previous studies have shown that metacognitive abilities are clearly linked to student success in terms of understanding what strategy to apply to solve a problem, and that this can be scaffolded in some situations to allow the learner to solve problems he otherwise would not be able to.

While metacognitive ability seems to be strongly linked to problem solving ability, it appears that metacognitive knowledge and intellectual aptitude are unrelated. (Sperling et. al., 2000). In other words, the amount of content knowledge you have and how you choose to use that knowledge in school is not directly related to your level of metacognition.

Swanson showed this to be the case when she asked 5th and 6th graders of high and average intellectual ability to solve pendulum and fluid combination problems after first completing a self-report interview designed to measure metacognitive knowledge about self, tasks, and strategies. He found that high metacognition/low aptitude students solved age-appropriate problems faster than low metacognition/high aptitude students, even though the two did not differ with respect to strategy used. This suggests that metacognitive knowledge can compensate for performance deficits (Swanson, 1990).

Desoete and her colleagues conducted a study with elementary school children aimed at contributing data to the debate on whether there are two or three components within metacognition. They said that, “in order to do so, we investigate empirically in two exploratory studies whether some of the most used metacognitive parameters (declarative knowledge, conditional knowledge, procedural knowledge, prediction, planning, monitoring, evaluation and attribution) can be combined into two (knowledge and skills) or three (add beliefs) supervariables on which young children differ” (Desoete, Roeyers, & Buysse, 2001, p. 436).

A post-hoc, follow-up analysis revealed the same findings as seen in the Swanson study. Above-average performers scored higher than average and below-average performers on measures of global metacognition. “Both studies have shown metacognition to be characteristic for the above-average (expert) approach to mathematical problem solving in the elementary school” (Desoete et al., 2001, p. 445). Although there may not be studies that show a correlation between metacognition and aptitude, clearly there is a correlation between metacognition and problem solving.

The previous studies employed elementary school students as subjects. Sperling and her colleagues were interested in conducting a study that used preschoolers as the subjects since, “little previous work has assessed preschoolers’ metacognitive regulation across multiple tasks” (Sperling et al., 2000, p. 235). They also say that most previous work has lacked assessment of individual’s metacognition across domains because “research investigating metacognitive regulation in academic domains has most often addressed learners’ metacognition while reading or problem-solving” and that both tasks require advanced skills which are harder to measure in young learners (Sperling et al., 2000, p. 235).

“The study was designed to address new aspects of self-regulation and to investigate connections between self-regulation as measured by theory of mind tasks and metacognitive regulation on age-appropriate problem-solving tasks” (Sperling et al., 2000, p. 239). There were 39 preschool children, 17 boys and 22 girls, from two child-care settings who participated in the study. Most of the children enrolled in these two preschools have parents who are affiliated with a land-grant university or accompanying medical center. At the time of the study, all the children were between 36 and 68 months old, with the 48th month used to split the group into two. All the children spoke English as their first language and gender and age were used as the independent variables. The dependent measures used were: performance on tasks, predictions and post-predictions of performance, and strategies used (Sperling et al., 2000). They found that even at such a young age, the children were “relatively effective at monitoring their own learning” (Sperling et al., 2000, p. 247). They also found that there was a strong relationship between theory of the mind (a metacognitive construct) and strategy use. It appears that even with preschoolers, those who are more metacognitively aware are better able to problem solve by selecting the appropriate strategy to use.

ASSESSMENT TOOL LITERATURE

Inventory of Metacognitive Self-Regulation

One of the problems with metacognitive studies that involve children is that there has not been a valid inventory of metacognition that has been reliability tested with large groups of students.

In the original study, the ratio of the students prediction score over the actual score was used as a way of determining a student's metacognitive awareness level. The belief was that the better a student was at predicting his score, the greater his level of metacognitive awareness. This method, called a calibration technique, is sometime used because it is easy to administer and score (Sperling, et. al., 2002). Sperling and colleagues also say that one concern with calibration techniques is that they may capture "self-efficacy or other motivational constructs rather than metacognition" (Sperling, et. al., 2002, p. 54). Pajares agrees saying that students' academic performances are partly the result of what they come to think they can accomplish and that is why "students' academic performances may differ markedly when they have similar abilities" (2002, p. 116).

The Inventory of Metacognitive Self-Regulation (IMSR) was designed in 2000 to be an easy-to-use, self-report inventory for use with children that focuses on metacognitive awareness and regulatory skills. Face validity was an important goal to the authors of the IMSR because they wanted an inventory that could be widely used by classroom teachers (Howard et. al., 2000a). Since that time reliability analyses and item factor analysis have been conducted numerous times, including three times with large populations of students, (n=829, n=1163, and n=1502), from all across the United States. The IMSR has a reliability alpha of .935.

The main author of the IMSR, Howard, had helped to develop the Junior Metacognitive Awareness Inventory (JrMAI) (Sperling et. al., 2002) that was designed to measure metacognitive awareness across all subject areas. While the results from validation analyses of the JrMAI indicate that it is a sound measure with reliable items (Sperling et. al., 2002), it does not demonstrate multiple factors, (Knowledge of Cognition and Regulation of Cognition), and a one factor interpretation is the best.

In comparison, the IMSR demonstrates that not only are there five particular metacognitive and self-regulatory constructs (listed in Table 3) relevant to problem-solving,

but that these constructs are “independent and therefore a student may show preferences or styles of metacognitive strength and weaknesses that depend upon his or her unique combination of constructs” (Howard et al., 2000b, p.5).

Motivation Scale: Student Opinion Scale

“When test results are analyzed, score interpretations are often questioned on several fronts: it is possible that a lack of motivation to perform well on these tests may produce scores that are spuriously low” (Sundre, 2000, n.p.). A motivational scale might be of use in informing a researcher of how much effort a student gave when completing an assessment. In 1999, Sundre revised a motivational scale published by Wolf and Smith in 1995. While Wolf and Smith’s scale was considered unidimensional, the scale revised by Sundre added two items and modified the wording of other items in order to delineate two factors (Sundre, 2000, n.p.).

Table 3. IMSR Constructs

Constructs	Definitions
Knowledge of Cognition	Understanding the extent and utilization of one’s unique cognitive abilities and the ways one learns best.
Objectivity	Standing outside oneself and thinking about one’s learning as it proceeds. (Self-regulatory construct pertinent to problem solving).
Problem Representation	Understanding the problem fully before proceeding. (Self-regulatory construct pertinent to problem solving).
Subtask Monitoring	Breaking the problem down into subtasks and monitoring the completion of each task. (Regulation of cognition factor).
Evaluation	Double-checking throughout the entire problem-solving process to evaluate if it is being done correctly. (Regulation of cognition factor).

The Student Opinion Scale (SOS) is a ten-item, five-point scale. Summing the responses to all ten items provides the Total Motivation score. The sum of responses to questions 1, 3, 4, 5, and 8 provides the Importance scale, while the responses to the other questions, 2, 6, 7, 9, and 10, presents the Effort scale. The Importance scale “provides a

measure of the personal relevance of the test to the examinee,” while the Effort scale “measures the level of effort students engaged in during the assessment task” (Sundre, 2000, n.p.).

SUMMARY

Although researchers have been continuously studying metacognition since the early studies of Flavell and Brown, clearly there is even now much that is not known and researchers today are still struggling with a comprehensive definition of metacognition. Although fewer metacognitive studies employ children as subjects, perhaps finding adults easier to study and gain access to, recent metacognitive studies with children demonstrate the following: (a) There is a positive relationship between teaching students metacognitive strategies and their ability to then choose the appropriate strategy to solve problems. (b) There is a positive relationship between IQ and metacognitive ability. (c) Although one might think that since there exists a positive relationship between IQ and metacognition, there would be a positive relationship between metacognition and student success in school on academic performance measures, that in fact, no such relationship appears to exist, although there is a positive relationship between metacognition and problem solving (most likely because of appropriate strategy use). Upon further analysis, this apparent lack of relationship is not so surprising, given that just because a child (or adult) has the ability to do well does not mean that he has the desire to do so. The study examined this interaction between metacognitive ability to perform well on posttests and student incentive on those same posttests.

CHAPTER THREE

METHODOLOGY

Chapter three begins with a review of the research questions and hypotheses. It also presents the context of the study, a description of the population and sample, instrumentation, data collection procedures and data analysis.

REVIEW OF THE RESEARCH QUESTIONS

As seen previously in Table 1, the study featured two independent variables: metacognitive prompts (present or absent) and incentive (present or absent).

In addition, there were several dependent variables: student results on the IMSR, posttest scores, metacognitive self-assessment probe answer, and gender.

Listed in Table 4 is a summary of the research questions, hypotheses and methods summary.

CONTEXT OF THE STUDY AND DESCRIPTION OF STUDENT POPULATIONS

The City of Chula Vista, with a population of over 200,000, is the eighth fastest growing city in the country and the second largest city in San Diego County (Clark, 2004, n.p.). Chula Vista expects to increase its population by 55,000 between 2000-2007, with 97% of the growth expected to take place on the city's east side (Moran, 2004, n.p.). San Diego Union-Tribune staff writer Chris Moran notes that, "Interstate 805 is Chula Vista's Mason-Dixon line, the proverbial 'tracks'" (n.p.). East of the 805, there are new homes with landscaped roads and new schools. On the west side, long-established neighborhoods, complete with heavy industry, (Moran, 2004). These city differences are apparent in the schools. Two public school districts serve Chula Vista K-12 students—the Chula Vista Elementary School District and the Sweetwater Union High School District. In both districts, schools west of the 805 normally have "higher percentages of Latinos, low-income and non-English-speaking students and lower state academic rankings" (Moran, 2004, n.p.).

Table 4. Question Summary

Research Question	Hypotheses	Operationalized with Method Summary
Main Question: Do metacognitive prompts increase student awareness of their own knowledge?	Prompted students will better predict their actual scores on posttest measures of comprehension than learners in the control group. The null hypothesis is that there will not be a significant difference between the experimental and control groups.	Multiple linear regression calculated with prediction scores and prompt status as the independent variables and posttest score as the dependent variable.
Question 2: Do metacognitive prompts increase student performance of posttest measures of recall of tutorial content? (Replication)	Experimental and control groups receiving treatments similar to those employed in the Watson study will yield similar results: no significant difference between prompted and non-prompted students. (See Table 2 for predicted rank order of mean scores).	Univariate test to analyze posttest scores (dependent variable) by group (fixed factor), with appropriate post-hoc tests.
Question 3: Is gender a factor in the effect of prompts on posttest performance? (Extension)	Posttest scores for boys promised rewards for will not differ significantly from girls promised similar rewards, whereas scores for boys not promised rewards will be significantly lower than girls not promised rewards. (See Table 2 for predicted rank order of mean scores).	Univariate test to analyze posttest scores (dependent variable) by group and gender (fixed factors), with appropriate post-hoc tests.

Table 4. (continued)

Research Question	Hypotheses	Operationalized with Method Summary
Question 4: What differences do boys and girls, both prompted and non-prompted, report about their behaviors and thinking regarding review of tutorial content following announcement of the posttest?	Boys who receive an external motivator to perform well on the posttest measures, will report staying focused on maximizing their posttest performance, while boys who do not receive an external motivator will report greater exploration of the navigational controls. Girls will not report any difference in behavior, whether or not they receive an external motivator.	Qualitative: Conduct interviews to learn first-hand how students reacted to metacognitive prompts. Quantitative: Independent sample t test (variables incentive and Student Opinion Scale) and Univariate test to analyze Student Opinion Scale scores (dependent variable) by group (fixed factor), with appropriate post-hoc tests.
Question 5: What is the relationship between prior metacognitive ability and the effectiveness of metacognitive prompts?	The underlying assumption is that students with high levels of metacognitive ability or aptitude are already able to employ their metacognitive abilities to study materials and prepare to answer questions, and therefore are unlikely to benefit as much from prompting as are students with lower metacognitive abilities. An exploratory analysis will be conducted, to determine if such a relationship exists.	IMSR used as a covariate. Exploratory analysis conducted to determine whether and to what extent factors based on these subscales might account for variance in the measures associated with Research Questions 1-4, considered as dependent measures.

Established in 1892, the Chula Vista Elementary School District (CVESD) is the largest K-6 grade school district in California with 2,650 staff serving a student population of about 25,600. The district opened its 42nd school in September 2004 (Chula Vista Elementary School District, 2004, n.p.). Data were collected at Casillas, Hilltop Drive, and Parkview Elementary schools in May-June of 2004.

The site from the original study, Joseph Casillas Elementary School is located off East J Street in an area of Chula Vista commonly referred to as Rancho Del Rey. The school, which opened in 1998, employs 42 full-time credentialed teachers and serves 823 students in grades K-6.

Led by principal, Dr. John Nelson, Casillas Elementary is on a single-track year-round schedule. The student body is ethnically diverse as illustrated in Table 5.

Table 5. Casillas Student Ethnicity

Racial/Ethnic Categories	Number of Students	Percentage of Students
African-American	43	5.20
American Indian	4	.05
Asian-America	69	8.40
Filipino-American	121	14.70
Hispanic or Latino	375	45.60
Pacific Islander	5	.06
White (Not Hispanic)	206	25.00

Casillas students perform well on standardized tests, with 46% of the 5th grade students scoring at the Advanced or Proficient levels on the 2001 California Standards Tests (CST); statewide, only 28% of 5th graders performed at this level. The California Standards Tests were developed specifically to assess students' performance on California's Academic Content Standards (Harcourt, 2004). In addition, Casillas' 2000-2001 API score was 789, narrowly missing the State of California's goal of 800. API stands for Academic Performance Index and is a school performance measurement system that was first developed as part of California's 1999 Public Schools Accountability Act. The Stanford 9 (reading and mathematics) Test, and the California Standards Test were used to calculate the 2000-2001 API scores.

A smaller and older school, Hilltop Drive Elementary opened in 1947 and is on a traditional school calendar. Principal Bruce Ferguson leads 30 teachers in serving 559 students in grades K-6. Hilltop Drive is located less than a mile west of Interstate 805 in an area of Chula Vista referred to as Hilltop.

As depicted in Table 6, over 60% of the Hilltop student body is Hispanic with White non-Hispanics constituting the second largest ethnic group.

Table 6. Hilltop Drive Student Ethnicity

Racial/Ethnic Categories	Number of Students	Percentage of Students
African-American	14	2.5
American Indian	6	1.1
Asian-America	6	1.1
Filipino-American	15	2.7
Hispanic or Latino	347	62.1
Pacific Islander	3	.05
White (Not Hispanic)	168	30.1

Hilltop Drive students also performed well on the 2001 California Standards Test with 41% of the 5th grade students scoring at the Advanced or Proficient levels. In addition, Hilltop Drive's 2000-2001 API score was 751, with a similar school rank of 10.

Parkview Elementary opened in 1969 and is just east of 805 South Chula Vista. Principal Charles Padilla and 24 teachers work at the smallest traditional school in the district, to educate the 441 students in grades K-6. As shown in Table 7, almost 60% of the student body is Hispanic, with White Non-Hispanics and Filipino-Americans constituting the next two largest ethnic groups.

While east of the 805, Parkview Elementary is in an area of Chula Vista where schools normally do not perform as well on standardized tests as the rest of the schools that are further east of 805. Still, Parkview students performed almost as well on the 2001 California Standards Test as Hilltop Drive with 37% of the 5th grade students scoring at the Advanced or Proficient levels. In addition, Parkview's 2000-2001 API score was 765, with a similar school rank of 10.

Table 7. Parkview Student Ethnicity

Racial/Ethnic Categories	Number of Students	Percentage of Students
African-American	17	3.9
American Indian	3	.7
Asian-America	13	2.9
Filipino-American	56	12.7
Hispanic or Latino	260	59.0
Pacific Islander	0	.0
White (Not Hispanic)	92	20.9

Table 8 shows a side-by-side comparison of the schools visited for data collection.

Table 8. School Comparisons

	Casillas	Hilltop Drive	Parkview
Year school opened	1998	1947	1969
Number of students	823	559	441
Number of teachers	42	30	24
Percentage of students scoring at advanced or proficient level on CST (2001)	46	41	37
API score (2001)	789	751	765
Similar school rank (2001)	N/A	10	10

POPULATION SAMPLE

All study the subjects were fifth graders, ages 10-12, at Casillas, Hilltop Drive, or Parkview Elementary. The principal investigator hoped to recruit approximately 140 students, about 50% more than Watson's original subject pool (students returned 171 positively signed parent permission slips). Following the same procedure used in Watson's study, English Language Learner (ELL) students were excluded from participation unless they were at or above an intermediate level, that is, unless they scored at least a three on the oral, and at least a two on the reading and writing portion of the California English Language Development Test (CELDT) (CTB McGraw-Hill, 2004). In the end, the researcher dropped no students from the analysis solely because of ELL status.

INSTRUMENTATION

Online Tutorial

Watson's (2001) original online tutorial consisted of 34 screens that included two short on-line posttests. An assistant recreated the original tutorial and an additional eight questions added—bringing the number of screens to 42. Students move through the tutorial sequentially for an introduction to basic concepts related to weather phenomena, such as clouds. The content for much of original tutorial was adapted, with permission, from materials created by the National Oceanic and Atmospheric Administration (NOAA) by an instructional designer with elementary school science teaching experience. In addition to adding posttest questions, the “play” button found in the original study was eliminated from the updated tutorial. In Watson's original study, each screen included a “play” button that the student could press to hear the words on the screen read aloud. Watson wanted to know if students who used the “play” feature would score differently on the posttests than those who did not use the feature. Less than five students pressed the play button more than once, a number not high enough to allow for significant statistical analysis.

The experimental groups completed a version of the tutorial that had a metacognitive prompt directly before each of the two posttest sections of the lesson. The control group did not see the metacognitive prompts, but instead saw pre-test screens saying, “Next: A Quiz.” Both groups saw a metacognitive probe question just before the start of the first and second posttests, as well as one at the very end of the lesson. These probes asked the students how well they thought they would perform/have performed on the posttests.

The modified lesson primarily emphasizes sequential navigation, which according to Davidson-Shivers, Shorter, Jordan, and Rasmussen (1999), is the type of navigation that fifth grade students frequently choose when they when using multimedia to prepare for testing situations. The tutorial screens contain buttons that allow the students to move forward and backward through the tutorial. In addition, a button to a “Where Am I?” diagram allows subjects to view where they are located and jump to any point in the lesson. The students control the pacing of the lesson.

Tutorial Posttests

As mentioned previously, eight new posttest questions were included in the recreated tutorial—four on the first posttest dealing with forms of water and four on the last posttest dealing with types of clouds. The researcher hypothesized that increasing the number of posttest questions would increase the reliability of the posttest, although the original study posttest reliability of .67 was quite good.

Inventory of Metacognitive Self Regulation

In addition to completing the posttests in the online tutorial, students also completed the Inventory of Metacognitive Self Regulation (IMSR). Howard et. al. (2000a) designed the IMSR as an easy-to-use, self-report inventory for use with children. The IMSR focuses on metacognitive awareness and regulatory skills. Several item reliability and factor analyses, including three studies with relatively large populations ($n=829$, $n=1163$, and $n=1502$) of US students, attest to the reliability of the IMSR. Current statistics suggest the IMSR has a reliability alpha of .935.

The main author of the IMSR, Howard, previously helped develop the Junior Metacognitive Awareness Inventory (Jr.MAI)—designed to measure metacognitive awareness across all subject areas. While the results from validation analyses of the Jr.MAI indicate that it is a sound measure with reliable items, it does not demonstrate multiple factors (Knowledge of Cognition and Regulation of Cognition), and a one-factor interpretation is the best (Sperling et. al., 2002). While the Jr.MAI has been tested and validated several times, because only a one-factor interpretation is best (independent constructs cannot be used), it is not as robust as the IMSR and does not allow a researcher to address data related to specific constructs or components of metacognition.

In comparison, the IMSR creators assert that not only are there five particular metacognitive and self-regulatory constructs (listed in Table 3) relevant to problem-solving, but that these constructs are “independent and therefore a student may show preferences or styles of metacognitive strength and weaknesses that depend upon his or her unique combination of constructs” (Howard, McGee, Shia, & Hong, 2001b, p. 4).

SUBJECT RECRUITMENT AND RANDOM ASSIGNMENT

The researcher dropped off parent-permission packets for all fifth grade students, as well as small incentives (such as pencils and bookmarks) at each of the three school sites. Students took home the packets. Although the researcher told parents not to return the permission form if they chose to not have their children participate in the study, some opted to return it anyways—accompanied a note requesting exclusion. The researcher stopped by each school several times to pick up returned packets from the teachers.

After collecting the parent letters, students were divided by gender and then assigned, by the researcher's assistant, to one of four groups using a random number generator to ensure random selection. The researcher's assistant then created log-in sheets for each student which included not only the student's log-in number and but also a note that reminded all students this was an important scientific experiment. The note for students in the incentive group also said, "For every question you get right on the quizzes, you will get a nickel. Wendy will return in a week to let you know how you did on the quizzes" (See Appendix H).

CHAPTER 4

RESULTS

In this study, the researcher examined the effect of metacognitive prompts on learner interactions as they completed an online tutorial about weather and cloud formation. The online tutorial was 42 screens in length and explored states of matter and cloud formations. The research assistant divided students by gender and then randomly assigned them to the control group or one of three treatment groups (Table 1). Groups 1 and 3 comprise the replication portion of the study, and groups 2 and 4 the extension.

This chapter describes the subjects used in the study, reviews the data collection procedures, explains how test reliability was obtained, and presents the findings (organized by research question).

SUBJECTS

The study population consisted of 147 fifth-grade students who attended Casillas, Hilltop Drive or Parkview Elementary School in the Chula Vista Elementary School District. The number of classes participating at each school site varied according to student enrollment. Five classrooms participated from Casillas, the largest school in the study, while two classrooms participated from Parkview, the smallest school in the study. Hilltop Drive had four classrooms participate. All fifth-grade students were given 11 x 14 sized envelopes to take home to their parents. The envelopes contained: (a) two parent permission slips, one to return and one to keep as a copy; (b) one student permission slip for the parent to keep as a copy, and (c) a business sized envelope, addressed to "5th Grade Teacher," for returning the signed parent permission slip. Of the 282 envelopes sent home, 171 were returned with a positive endorsement of a student's participation in the study. The return rate of 60% almost mirrors the original study return rate of 63%. The researcher gave the teachers at each school site small incentives (pens, bookmarks and key chains) to give to students as they returned the forms. Teachers were diligent about reminding students to bring back the permission slips.

The researcher decided to test all students whose parents agreed to their participation (and who themselves agreed to participate), and then remove the scores of any student who did not meet the literacy requirement. Two students opted out of the study whose scores would have been dropped, and a few students who were ELL classified failed to answer any questions on the tutorial. No scores were removed solely because of the literacy requirement. This is probably partly due to the fact that the researcher did not send any forms home in Spanish (and therefore the return rate for parents who speak only Spanish was probably lower).

POSTTEST INSTRUMENT RELIABILITY AND VALIDITY

In his original study, Watson (2001) ran a split-half reliability test on the ten-item posttest measure. Watson found that the total posttest reliability had a value of 0.67.

In this replication study, the researcher added eight questions (four to each posttest) in hope of increasing posttest reliability. Calculating Cronbach's alpha, the reliability (internal consistency) of the posttest measure was $\alpha = .62$. The original ten-item posttest questions were also analyzed for reliability using Cronbach's alpha and had a level of $\alpha = .50$. Apparently in the replication study, if only the original ten-item questions had been used the reliability would not have been as high as when the additional eight items were added. Still, $\alpha = .62$ is a lower level of reliability than in the original study, $\alpha = .67$.

DATA COLLECTION

Online tutorial data were collected in the 2003-2004 school year from Casillas Elementary in the afternoons between May 24th and 27th, at Hilltop Drive in the afternoons between May 28th and June 3rd, and at Parkview on June 10th. The Inventory of Metacognitive Self-Regulation (IMSR) was administered to all fifth-grade students, in their classrooms as a group, about a week prior to visiting each school for the online tutorial data collection.

The data collection process was the same each day. By teacher preference, the students were called out of class using a call-slip and reported to the researcher at the library (Casillas and Hilltop Drive) or empty classroom (Parkview).

Upon entering the library or classroom, the students listened to an orienting movie on a laptop computer, (see Appendix G: Instructions for Subjects) and then reviewed and signed the student consent form. After seeing the orienting movie and reading the student consent form, four students choose not to participate in the study and one student asked to be called from her class the following day because she was worried about completing an assignment that was due. Two of the students did not provide a reason for opting out, and two explained that their language skills were not “good enough” to participate.

Students who signed the student consent form were given their login paper and asked to read the letter on it (see Appendix H) that stated the following: “This is a very important scientific experiment. Please help me by not telling anyone about what you read on this paper, or about the computer tutorial.” Students receiving an incentive also read, “For every question you get right on the quizzes, you will get a nickel. Wendy will return next week to let you know how you did on the quizzes.” The student login papers contained not only the letters, but also a pre-assigned (by random numbering) login number for each student. After student read the letters, they were logged on to a computer by the researcher or her assistant.

As mentioned previously, there were two versions of the tutorial, both with the same number of screens. Groups Three and Four saw a metacognitive prompt directly before each posttest (see Table 1), instead of just a screen announcing that a test was next. User data were automatically collected in an online database as students completed the tutorial, thereby eliminating manual data entry.

Most students took 15-20 minutes to complete the tutorial, with one student taking 45 minutes. No student chose to stop partway though the tutorial, although six did not answer any of the questions. In total, 167 students completed the online tutorial (83 males and 64 females). The researcher dropped 14 students because they did not correctly answer the “check for reading” question on the IMRS. Of these students, three were from Casillas (all male), six were from Hilltop Drive (three male and three female) and five were from Parkview (one male and two female). Of the six students who were dropped because they did not answer any of the questions, four were from Casillas (two males and two female) and two were from Parkview (both males).

Students completed the online tutorial and then the Student Opinion Scale. Interviews followed and lasted between five and 10 minutes. Students were thanked and sent back to

class with a call slip for the next student. This process continued until all research subjects had completed the online tutorial.

Eight students, all boys, closed the software application accidentally while using the tutorial. In each case, the researcher restarted the application and the student was told to navigate to where he left off.

Six students were confused when they arrived at the 4th screen of the lesson which poses the questions: “What are clouds? Where do they come from?” As with the original study, these students thought the screen was a posttest item and didn’t know how to answer. In each case, the subject asked the researcher for help. Following the example from the original study, the researcher said, “You don’t have to answer these questions. The lesson is just saying that these are the questions that will be answered during the lesson.” The students then proceeded to click the forward button without any additional prompting.

Once all data had been collected from a school site, the researcher took a thank-you letter to each student with a dollar bill enclosed inside. Some students (two or three at each site) in the incentive group were surprised that the researcher returned with the promised reward. Students in the groups that had not been promised an incentive were very happy to receive the dollar.

RESEARCH QUESTION ONE

The main research question asked, “Do metacognitive prompts increase students’ awareness of their own knowledge?” The researcher hypothesized that prompted students would be better at predicting their actual scores on posttest measures of comprehension than learners in the control group because they would have been prompted by the metacognitive prompt to think about what they did and did not know.

To test this hypothesis, two multiple linear regressions models were estimated (one for each posttest) with prediction scores and prompt status as the independent variables and posttest scores as the dependent variable. For the first posttest—the results of which are depicted in Table 9—the regression equation was significant ($F(2, 142) = 4.473, p = .005; R^2 = .086$). However, it does not test the hypothesis, the interaction variable (Prompt * Prediction) does. The only significant variable (Table 10) is for Prediction ($B = .83; p = .000$). Most of the variance noted in Table 9 is accounted for by the main effect of Prediction

(see Table 10). So, while students in general did a fair job of predicting how well they would do on the posttest, prompted students did not predict any more accurately than non-prompted students. This is consistent with Schneider and Pressley's findings (1997) that indicate prediction abilities in children show a strong tendency towards improvement at the 4th and 5th grade level.

Table 9: Model Summary for Posttest One

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.292 ^a	.085	.072	1.476
2	.293 ^b	.086	.067	1.481

^a Predictors: (Constant), Prompt, Prediction Test One

^b Predictors: (Constant), Prompt, Prediction Test One, Interaction

Table 10: Regression Analysis Summary for Variables Predicting the Interaction Effect Between Students' Score Prediction on Posttest One and Prompted Status (N = 146)

Variable	B	SE B	β
Step 1			
Constant	3.76	0.57	
Prediction	0.83	0.23	0.29*
Prompt	0.13	0.24	0.04
Step 2			
Constant	3.94	0.83	
Prediction	0.75	0.34	0.26
Prompt	-0.20	1.11	-0.07
Interaction	0.14	0.46	0.11

Note. $R^2 = .085$ for Step 1; $R^2 = .086$ for Step 2.

* $p = .000$

The regression equation was not significant for the second posttest either ($F(2, 142) = 5.787, p = .270; R^2 = .027$) (see Table 11). In addition, none of the main effects, nor the interaction effect was significant (see Table 12). The students in the prompted group did not predict more accurately how well they would do on the second posttest (the same findings as in posttest one), and in addition, none of the students predicted very accurately how well they would do on the second posttest (unlike the results seen in posttest one).

Table 11: Model Summary for Posttest Two

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.15 ^a	.02	.008	2.09
2	.16 ^b	.03	.007	2.09

^a Predictors: (Constant), Prompt, Prediction Test Two

^b Predictors: (Constant), Prompt, Prediction Test Two, Interaction

Table 12: Regression Analysis Summary for Variables Predicting the Interaction Effect Between Students' Prediction of Scores on Posttest Two and Prompted Status (N = 146)

Variable	B	SE B	β
Step 1			
Constant	3.67	0.77	
Prediction	0.16	0.09	0.15
Prompt	0.12	0.35	0.03
Step 2			
Constant	2.92	1.12	
Prediction	0.26	0.14	0.23
Prompt	1.44	1.47	0.34
Interaction	-.17	0.18	-0.33

Note. $R^2 = .021$ for Step 1; $R^2 = .027$ for Step 2.

Research Questions Two and Three

The second research question asked, "Do metacognitive prompts increase student performance of posttest measures of recall of tutorial content?" The researcher hypothesized that experimental and control groups receiving treatments similar to those employed in the Watson (2001) study would yield similar results: no significant difference between prompted and non-prompted students, *until* the results were analyzed by gender.

In Watson's study (2001), posttest scores analyzed by gender (in a post hoc test) showed that while the main effects of gender and prompted condition were not significant, the interaction effect between gender and prompted condition was significant at the .001 level. Prompted girls had a mean score ($m = 5.65$) higher than non-prompted girls ($m = 3.62$), while prompted boys ($m = 3.67$) did not have a higher mean score than non-prompted boys ($m = 5.2$). Watson (2001) suggested that boys and girls might be interpreting

the metacognitive prompt differently, and therefore reacted differently to it (girls stayed focused on the task of performing well on the posttests and boys explored the online tutorial more because they did not have an incentive to perform well on the posttests). Because of the expected gender difference, question three asked, “Is gender a factor in the effect of prompts on posttest performance?”

Because of Watson’s findings (2001), the researcher hypothesized that in the replication and extension study the posttest scores for boys promised rewards would not differ significantly from girls promised similar rewards (the external incentive would keep boys on task), whereas scores for boys not promised rewards would be lower than girls not promised rewards (as in Watson’s study). See Table 2 for a review of the predicted scores.

A factorial ANOVA was used for analysis of data in both questions two and three. For question two, a 2 x 2 between-subjects factorial ANOVA was calculated comparing the final exam scores by prompted condition and gender. A main effect for prompt was not significant ($F(1, 143) = .113, p = .74$). The main effect for gender was also not significant ($F(1, 143) = 1.31, p = .26$). Finally, the interaction was not significant ($F(1, 143) = .000, p = .99$). Thus, as expected, neither gender nor metacognitive prompting significantly affected posttest scores. It was hypothesized though that the interaction effect would be significant but it was not.

For question three, a 2 x 2 x 2 between-subjects factorial ANOVA was calculated comparing the final exam scores by prompted condition, incentive condition, and gender. A main effect for prompt was not significant ($F(1, 139) = .04, p = .84$). The main effect for gender was not significant ($F(1, 139) = 1.22, p = .27$). The main effect for incentive was not significant ($F(1, 139) = .53, p = .47$). In addition, none of the interaction effects were significant, except for the one between gender, prompt and incentive ($F(1, 139) = 5.45, p = .02$) (see Tables 13 and 14).

Research Question Four

Research Question Four was, “Are there differences reported by boys and girls, both prompted and non-prompted, about their behaviors and thinking regarding review of tutorial content following announcement of the posttest?”

Table 13. Descriptive Statistics for Question Three

	Gender	Incentive	Mean	Std. Deviation	N
Prompt	Male	Incentive	11.19	2.27	21
		No Incentive	11.00	3.45	20
		Total	11.10	2.87	41
	Female	Incentive	10.13	3.03	16
		No Incentive	10.89	3.03	18
		Total	10.53	3.01	34
	Total	Incentive	10.73	2.64	37
		No Incentive	10.95	3.21	38
		Total	10.84	2.92	75
No Prompt	Male	Incentive	10.55	3.16	22
		No Incentive	11.35	2.54	20
		Total	10.93	2.87	42
	Female	Incentive	11.86	3.11	14
		No Incentive	9.06	2.67	16
		Total	10.37	3.17	30
	Total	Incentive	11.06	3.16	36
		No Incentive	10.33	2.81	36
		Total	10.69	2.99	72
Total	Male	Incentive	10.86	2.75	43
		No Incentive	11.18	2.99	40
		Total	11.01	2.86	83
	Female	Incentive	10.93	3.14	30
		No Incentive	10.03	2.97	34
		Total	10.45	3.06	64
	Total	Incentive	10.89	2.89	73
		No Incentive	10.65	3.02	74
		Total	10.77	2.95	147

While Watson did not address this question with interviews or other self-reported data, he speculated that boys and girls might interpret metacognitive prompts differently, and therefore responded to them in dissimilar fashions. This speculation, that boys tend to

explore software programs more than girls and not stay as focused on the task, was consistent with studies conducted by Caftori (1996) and McKean, Hoffman and Allen (2000).

The hypotheses in the replication were:

- Boys who receive an incentive to perform well on the posttest measures, will report staying focused on maximizing their posttest performance, while boys who do not receive an incentive will report greater exploration of the navigational controls.
- Girls will not report any difference in behavior, whether or not they receive an incentive.

Table 14. Test of Between-Subjects Effects for Question Three

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	82.71 ^a	7	11.82	1.38	.22
Intercept	16625.59	1	16625.59	1946.19	.00
Prompt	.34	1	.34	.04	.84
Gender	10.49	1	10.41	1.22	.27
Incentive	4.51	1	4.51	.53	.47
Prompt * Gender	.09057	1	.09057	.01	.92
Prompt * Incentive	14.76	1	14.77	1.73	.19
Gender * Incentive	15.72	1	15.72	1.84	.18
Prompt * Gender * Incentive	46.59	1	46.59	5.45	.02
Error	1187.42	139	8.54		
Total	18317.00	147			
Corrected Total	1270.14	146			

^a R Squared = .07 (Adjusted R Squared = .02)

QUANTITATIVE ANALYSIS

An independent samples *t* test was performed with three Student Opinion Scale factors (Total Incentive, Importance Scale, and Effort Scale) serving as the dependent variables and incentive condition as the independent variable. The researcher found a significant difference between the means of the incentive and non-incentive students ($t(145) = -.80, p = .03$) on the Effort scale (Tables 15 and 16). The mean score of the non-incentive students ($m = 19.53, sd = 2.3$) was higher than the mean score of the incentive students ($m = 19.16, sd = 3.1$). No significant difference was found for either of the other

factors (Total Incentive or Importance Scale). While there is a significant statistical difference in mean score on the Effort Scale, it is not a difference that was hypothesized.

Table 15. Group Statistics

	Incentive	N	Mean	Std. Deviation	Std. Error Mean
Total Incentive	Incentive	73	37.22	5.45	.64
	No Incentive	74	37.64	4.37	.51
Importance Scale	Incentive	73	18.05	3.46	.40
	No Incentive	74	18.11	3.25	.38
Effort Scale	Incentive	73	19.16	3.11	.36
	No Incentive	74	19.53	2.34	.27

Table 16. Independent samples *t* test

		F	<i>p</i>	<i>t</i>	df
Total Incentive	Equal variances assumed	1.39	.24	-.51	145
	Equal variances not assumed			-.51	137.76
Importance Scale	Equal variances assumed	.30	.58	-.10	145
	Equal variances not assumed			-.10	144.14
Effort Scale	Equal variances assumed	4.58	.03	-.80	145
	Equal variances not assumed			-.80	133.65

A 3 x 2 x 2 factorial ANOVA was then calculated comparing the Student Opinion Scale factors (Total Incentive, Importance Scale, and Effort Scale) by prompted condition and gender. None of the main effects or interaction effects were significant (Tables 17, 18, and 19).

QUALITATIVE ANALYSIS

After completing the online tutorial and the Student Opinion Scale, a research assistant interviewed 69 (47%) students (39 male and 30 female; 32 incentive and 37 no incentive; 32 prompted and 37 not prompted).

Table 17. Tests of Between-Subjects Effects: Total Incentive

Source	Type III Sum of Squares	df	Mean Square	F	<i>p</i>
Corrected Model	158.20 ^a	7	22.60	.93	.49
Intercept	200681.49	1	200681.49	8248.48	.00
Incentive	9.24	1	9.24	.38	.54
Prompt	15.27	1	15.27	.63	.43
Gender	33.51	1	33.51	1.38	.24
Incentive * Prompt	15.08	1	15.08	.62	.43
Incentive * Gender	22.23	1	22.23	.91	.34
Prompt * Gender	.50	1	.50	.02	.89
Incentive * Prompt * Gender	71.82	1	71.82	2.95	.09
Error	3381.80	139			
Total	209472.00	147			
Corrected Total	3540.00	146			

^a R Squared = .05 (Adjusted R Squared = .003)

The majority 56 of the students (81%) reported not clicking the back button after seeing the prompt announcing that a test was next (Table 20). When asked during the interview why they had not clicked the back button before beginning the posttest, most responded with an answer that indicated they felt confident that they already knew the information needed to answer the upcoming posttest questions. Example responses included, “I was sure I knew the material,” “I had read it carefully,” and “I was ready for the test.”

Only 13 students (19%) reported clicking the back button directly after seeing the screen announcing that a posttest was next. All 13 reported doing so in order to review. Example responses included, “I wanted to see it again, to read the answers again,” “to go back and review,” and “I was not sure I had it all and I wanted to read it again.” While all thirteen gave similar verbal responses for clicking the back button, there appears to not be any other unifying factor among the students: eight prompted and five not prompted; nine girls and four boys; five incentive and eight not incentive.

In addition to asking the students if they clicked the back button and why or why not, the interviewer also showed the students a picture of the prompt (or control) screen and asked

them to recall, “What did you think about when you saw this screen?” The majority of the students simply repeated that they had known they were prepared for the posttest, but a few (seven) reported not thinking about anything, and two reported wondering how long the posttest would be. So, it appears that while there was no reported difference among the prompted and non-prompted girls (as expected), there was also no reported difference among the boys (unexpected).

Table 18. Tests of Between-Subjects Effects: Importance Scale

Source	Type III Sum of Squares	df	Mean Square	F	<i>p</i>
Corrected Model	47.74 ^a	7	22.60	.93	.49
Intercept	46894.58	1	46894.58	4116.98	.00
Incentive	.21	1	.21	.02	.89
Prompt	1.09	1	1.09	.10	.76
Gender	.72	1	.72	.06	.80
Incentive * Prompt	5.57	1	5.57	.49	.49
Incentive * Gender	6.79	1	6.79	.60	.44
Prompt * Gender	1.61	1	1.61	.14	.71
Incentive * Prompt * Gender	32.90	1	32.90	2.89	.09
Error	1583.28	139	11.39		
Total	49692.00	147			
Corrected Total	1631.02	146			

^a R Squared = .03 (Adjusted R Squared = -.02)

Research Question Five

Research Question Five asked, “What is the relationship between prior metacognitive ability and the effectiveness of metacognitive prompts?” Watson did not address this issue and did not identify instruments that might have been used to measure a child’s metacognitive abilities. Researchers (Allen & Merrill, 1985; Kurtz & Borkowski, 1987; Mevarech, 1999) suggest that metacognitive prompts are more likely to enhance performance of students who are unlikely to employ metacognitive skills on their own. The underlying assumption is that students with high levels of metacognitive ability or aptitude are already able to employ their metacognitive abilities to study materials and prepare to answer

questions, and therefore are unlikely to benefit as much from prompting as are students with lower metacognitive abilities.

Table 19. Tests of Between-Subjects Effects: Effort Scale

Source	Type III Sum of Squares	df	Mean Square	F	<i>p</i>
Corrected Model	69.89 ^a	7	9.98	1.35	.23
Intercept	53556.75	1	53556.75	7231.62	.00
Incentive	6.67	1	6.67	.90	.34
Prompt	24.54	1	24.54	3.31	.07
Gender	24.40	1	24.40	3.30	.07
Incentive * Prompt	2.32	1	2.32	.31	.58
Incentive * Gender	4.45	1	4.45	.60	.44
Prompt * Gender	3.89	1	3.89	.53	.47
Incentive * Prompt * Gender	7.50	1	7.50	1.01	.32
Error	1029.42	139	7.41		
Total	56122.00	147			
Corrected Total	1099.31	146			

^a R Squared = .06 (Adjusted R Squared = -.02)

Table 20. Crosstabulation

Prompt	Incentive			Male	Female	Total
No prompt	No incentive	Go back	Did not go back	12	7	19
			Went back	0	3	3
		Total		12	10	22
	Incentive	Go back	Did not go back	11	2	13
			Went back	2	0	2
		Total		13	2	15
Prompt	No incentive	Go back	Did not go back	5	5	10
			Went back	1	4	5
		Total		6	9	15
	Incentive	Go back	Did not go back	7	7	14
			Went back	1	2	3
		Total		8	9	17

As part of the extension study, students completed an inventory of metacognitive awareness and skills based on the work of Howard, McGee, Shia and Hong (2000a). Their inventory includes a number of sub-scales, including Knowledge of Cognition, Objectivity, Problem Representation, Subtask Monitoring and Evaluation. An exploratory analysis was conducted to determine whether and to what extent factors based on these subscales, considered as independent variables, might account for variance in the measures associated with Research Questions 1-4, considered as dependent measures.

IMSR

The researcher believed that there was a relationship between a student's metacognitive level (IMSR constructs) and his or her score on posttest measures. To examine a possible relationship, a Pearson correlation coefficient was calculated. A weak positive correlation (see Table 21) was found ($r(145) = .18, p = .03$ for construct four and a weak positive correlation was found ($r(145) = .20, p = .02$) was found for construct three.

A one-way between subjects ANCOVA was calculated to examine the effects of gender, incentive and a metacognitive prompt on posttest scores, covarying out the effect of the IMSR and the constructs.

As can be seen in Table 22, IMSR three and four were significantly related to posttest score, but the main effects for gender, prompt and incentive were not significant, even after covarying out the effect of the IMSR.

Table 21. Correlation of IMSR Constructs Three and Four to Posttest Scores

		IMSR 3	Total Test Score	IMSR 4
IMSR 3	Pearson Correlation	1	.20	.33
	Sig. (2-tailed)	.	.02	.00
	N	147	147	147
Total Test Score	Pearson Correlation	.197	1	.18
	Sig. (2-tailed)	.02	.	.03
	N	147	147	147
IMSR 4	Pearson Correlation	.33	.18	1
	Sig. (2-tailed)	.00	.03	.
	N	147	147	147

Table 22. ANCOVA

Source	Type III Sum of Squares	df	Mean Square	F	<i>p</i>
Corrected Model	204.26 ^a	9	22.70	2.92	.00
Intercept	252.26	1	252.26	32.42	.00
IMSR 3	82.46	1	82.46	10.60	.00
IMSR 4	79.33	1	79.33	10.20	.00
Gender	15.09	1	15.09	1.94	.16
Prompt	.76	1	.76	.10	.76
Incentive	1.31	1	1.31	.17	.68
Gender * Prompt	.14	1	.14	.02	.89
Gender * Incentive	10.03	1	10.03	1.29	.26
Prompt * Incentive	8.87	1	8.87	1.14	.28
Gender * Prompt * Incentive	38.23	1	38.23	4.91	.03
Error	1065.88	137			
Total	18317.00	147			
Corrected Total	1270.14	146			

^a R Squared = .05 (Adjusted R Squared = -.00)

Additional Data Analysis to Compare Posttests

DIFFERENCE BETWEEN THE TWO POSTTESTS

Since the researcher believed there might be difference between the two posttests, a paired-samples *t* test was calculated to compare the mean score of the first posttest to the mean score of the second posttest (both scores calculated as percentages). The percentage mean on the first posttest was .72 (*sd* = .19), and the percentage mean on the second posttest was .50 (*sd* = .21) (see Table 23). A significant difference between the two posttests was found ($t(146) = 11.43, p = .00$) (see Tables 24 and 25).

Table 23. Paired Samples Statistics for Posttests One and Two

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Percent Test One	.72	147	.19	.02
	Percent Test Two	.50	147	.21	.02

Table 24. Paired Samples Correlations for Posttests One and Two

		N	Correlation	<i>p</i>
Pair 1	Percent Test One & Percent Test Two	147	.30	.00

Table 25. Paired Samples Test for Posttests One and Two

		Paired Differences			<i>t</i>	df	<i>p</i> (2-tailed)
		Mean	Std. Deviation	Std. Error of Mean			
Pair 1	Percent Test One & Percent Test Two	.22	.24	.02	11.43	146	.00

ANALYSIS FOR POSTTEST ONE ONLY

Upon realizing that there was a significant difference between the two posttests, all the previous tests used in questions one through five were run, with the dependent variable being first posttest one and then posttest two (instead of the total posttest scores) to see if there were any differences. The only change from previous results found was that when a 2 x 2 between-subjects factorial ANOVA was calculated comparing the first posttest score by prompted condition; incentive condition; and gender, the main effect for gender was significant ($F(1, 139) = 4.83, p = .03$) (see Table 26). No other changes were found.

Table 26. Test of Between-Subjects Effects, Posttest One Only

Source	Type III Sum of Squares	df	Mean Square	F	<i>p</i>
Corrected Model	17.65 ^a	7	2.52	1.08	.38
Intercept	4754.79	1	4754.79	2031.21	.00
Gender	11.31	1	11.31	4.83	.03
Prompt	.45	1	.45	.19	.66
Incentive	4.657E-03	1	4.657E-03	.00	.96
Gender * Prompt	.54	1	.54	.23	.63
Gender * Incentive	2.76	1	2.76	1.18	.28
Prompt * Incentive	5.170E-02	1	5.170E-02	.022	.88
Gender * Prompt* Incentive	2.75	1	2.75	1.18	.28
Error	325.38	139	2.34		
Total	5258.00	147			
Corrected Total	343.03	146			

^a R Squared = .05 (Adjusted R Squared = .00)

CHAPTER 5

DISCUSSION

In *The Effect of Embedded Metacognitive Cues and Probes on Use of Learner Control Features in an On-line Lesson for Elementary Students* (2001), Watson found minimal prompting in an online tutorial increased 5th grade students' understanding of how much they understood. Prompted students predicted more accurately than non-prompted students their own performance on posttest measures, but did not score any greater. In a post-hoc analysis, Watson discovered gender differences. Prompted girls scored greater on posttest measures than non-prompted girls, while prompted boys scored lower than non-prompted boys. This study replicated and extended Watson's study. The researcher hoped to determine if the gender differences illuminated in the original study could be replicated, and if so, could those differences be explained.

MAIN RESEARCH QUESTIONS

The main research question asked, "Do metacognitive prompts increase student awareness of their own knowledge?" It was hypothesized that prompted students would be better at predicting their scores on posttest measures of comprehension than learners in the control group because they had been prompted to think about what they knew and didn't know by the metacognitive prompt.

The regression equation was significant for the first posttest (R^2 of .086), but further analysis noted the majority of this variance is accounted for by the main effect of prediction, and not by the interaction variable of prediction/prompt as was hypothesized. The data does not support the hypothesis, but instead reveals that prompted students predicted no more accurately than non-prompted students. However, unlike the original study, all students in the replication study predicted moderately well their scores on the first posttest ($B = .83$; $p = .000$).

While students were moderately accurate at predicting their own scores on the first posttest, this was not true with the second one. The researcher found a discrepancy in

students' prediction ability on the two posttests when analyzing the data. The prediction variable on the second posttest was much lower ($B = .16; p = .000$) than the prediction variable on the first posttest ($B = .83; p = .000$). The researcher expected to find that students predicted well on the second posttest, since they had on the first. This was not true and was an indicator that there might be differences between the two posttests.

RESEARCH QUESTIONS TWO AND THREE

While research question one examined if prompted students predict more accurately their scores on posttest measures, questions two and three posited whether they would perform better on those posttests than non-prompted students.

Research question two asked, "Do metacognitive prompts increase student performance of posttest measures of recall of tutorial content?" and question three asked, "Is gender a factor in the effect of prompts on posttest performance?"

With question two, the researcher hypothesized that prompted and non-prompted groups receiving treatments similar to those employed in the Watson study would yield similar results: no significant difference between prompted and non-prompted students, *until* the results were analyzed by gender.

In Watson's study, posttest scores examined by gender in post-hoc analysis showed that, while the main effects of gender and prompted condition were not significant, the interaction effect between gender and prompted condition was significant at the .001 level. Prompted girls had a higher mean score ($m = 5.65$) than non-prompted girls ($m = 3.62$), while prompted boys ($m = 3.67$) had a lower mean score than non-prompted boys ($m = 5.2$). Watson (2001) speculated that boys and girls might be interpreting the metacognitive prompt differently. Research (Caftori, 1996; McKean et. al, 2000) suggests that boys may be more likely to explore computer environments than girls and, in Watson's study it is possible that the girls were intrinsically motivated to stay focused on the task, while the boys instead chose to explore the online tutorial. Because Watson found gender differences, question three sought to determine whether those differences could be replicated.

As hypothesized, prompted students scored no greater on posttest measures than non-prompted students. The main effect for prompt was not significant ($F(1, 143) = .113, p = .74$) in the 2 x 2 between-subjects factorial ANOVA (incentive not used as an independent

variable), or in the 2 x 2 x 2 between-subjects factorial ANOVA ($F(1, 139) = .04, p = .84$) where incentive was included as an independent variable. The researcher expected the interaction effect for prompt and gender would be significant in the 2 x 2 ANOVA, as in Watson's study, but surprisingly it was not.

The only significant finding from the data analysis for questions two and three is difficult to explain. The researcher expected the interaction effect of gender and prompt to be significant, as this would have shown that boys and girls were interpreting the metacognitive prompt differently. Instead, in the 2 x 2 x 2 between-subjects factorial ANOVA the interaction effect of gender, prompt and incentive is significant, yet none on the main effects or other interaction effects are. Table 18 lists mean scores of each subject group. Although the mean scores are clustered around 10 and 11, there is a large gap noted with the scores of girls who were not prompted. Those girls who received an incentive scored higher ($m = 11.86$) than girls who did not get an incentive ($m = 9.06$). Instances such as this one led to the significant interaction effect.

RESEARCH QUESTION FOUR

Research question four asked, "Are there differences reported by boys and girls, both prompted and non-prompted, about their behaviors and thinking regarding review of tutorial content following announcement of the posttest?"

The hypotheses in the replication were:

- Boys who receive an incentive to perform well on the posttest measures, will report staying focused on maximizing their posttest performance, while boys who do not receive an incentive will report greater exploration of the navigational controls.
- Girls will not report any difference in behavior, whether or not they receive an incentive.

All students completed the Student Opinion Scale (SOS) after the online tutorial and 69 students (47%) were then interviewed. Interview data does not support the hypothesis about the boys. This is not surprising given the lack of significant findings with questions two and three. Boys who saw an incentive did not report staying on task more than boys who did not see an incentive. In addition, non-incentive boys did not report any greater exploration of the navigational controls than boys who saw an incentive. In fact, none of the boys in either group discussed the navigational controls at all with the interviewer, although

there was not a specific interview question about navigational controls. As stated before, eight boys closed the browser accidentally while using the tutorial. Even though the researcher noted each time this occurred, she did not write down the students' login numbers, and therefore does not know the boys subject groups.

The interviewer questioned students about clicking (or not clicking) the back button before taking the posttests, and most reported they already knew the material and wanted to continue. Only 19% of the students reported clicking the back button before the posttests. When closely examining responses of students who reported clicking back, the researcher found no common feature among the students other than their reporting wanting to review the material more before continuing with the posttest.

The researcher analyzed Student Opinion Scale data with an independent samples t test, followed by a 3 x 2 x 2 factorial ANOVA. The t test showed a significant difference ($t(145) = -.80, p = .03$) between the means of the incentive and non-incentive students on the Effort Scale, but not on the Importance Scale or on Total Incentive. This difference was not anticipated since the non-incentive students reported *more effort* than the incentive students. This result, along with the surprised reaction of some students when the researcher returned with the promised incentive, suggests that the incentive may not have been viewed as such by the students.

RESEARCH QUESTION FIVE

Research question five asked, "What is the relationship between prior metacognitive ability and the effectiveness of metacognitive prompts?" When a Pearson correlation coefficient was calculated for the relationship between the IMSR constructs and the posttest scores, a significant weak positive correlation was seen for constructs three (Objectivity: Standing outside oneself and thinking about one's learning as it proceeds) and four (Problem Representation: understanding the problem fully before proceeding). Since there were significant correlations, an ANCOVA was calculated to examine the effects of gender, incentive and prompt on posttest scores when covarying out the effects of IMSR constructs three and four. As seen in Table 28, using the IMSR as a covariate made no statistical difference when looking at the main effects. Although the IMSR is a well-validated and

reliable survey, it is possible that the subjects used in the study were at similar metacognitive levels, and therefore using the IMSR as a covariate was not effective.

STRENGTHS OF THE STUDY

While this study failed to replicate Watson's findings, several strengths are noted. First, as did Watson, this researcher used a true experimental design and took great care to avoid all possible bias. More research subjects were boys, a reflection of the schools' populations, and the researcher decided to use a stratified random sample to evenly distribute boys and girls among the research groups. Someone other than the researcher performed the random assignment, and the researcher was blind until after all subjects had completed the tutorial.

Another strength was the faithful replication of the original study, with extensions planned to not interfere with the replication. Although the original lesson was no longer available, screen shots had been saved and the tutorial was recreated using the exact graphics from the first lesson. Eight additional posttest questions were added, but at the end of the original questions. An additional metacognitive probe was included, but was added directly before the second posttest, near the end of the tutorial. An additional experimental group, student promised a small incentive to perform well on the posttest, was added, but in such a way that the original groups were not effected. All of these additions, needed for the extension, did not impede the replication.

The researcher was also faithful in subject recruitment. Casillas Elementary was the school used in the original study and Dr. Nelson, the principal since the school's opening in 1998, agreed to the replication study. Since additional students would be needed for the extensions, the researcher successfully sought to find schools that matched Casillas closely on the school accountability report cards.

Finally, the researcher added the Inventory of Metacognitive Self-Regulation (IMSR). Since the researcher also believed that the students' ability to predict their own scores was more a factor of self-efficacy, and not metacognition, she went looking for and found a well-validated metacognitive inventory designed to be used with children.

LIMITATIONS OF THE STUDY

While the replication study had several strengths, there were also some limitations. As in Watson's study, there was a reasonable attempt to control variance and experimenter bias by: randomly assigning students, showing each subject the same videotaped instructions, and using schools that performed similarly on standardized measures. Still, some of the unavoidable variance found in the first study is also present in this study. The researcher personally administered the online tutorial to each student, and although she tried to behave the same with each student, some unavoidable differences might have led subjects to react differently to the lesson. The lesson was administered in a library or spare classroom, and while the rooms were ideally suited to the situation, they were not the student's classrooms—the room a student would feel most comfortable in.

The researcher's concerns about the study can be gathered in the following three groupings: tutorial issues previously addressed by Watson, posttest limitations and differences, and thoughts regarding a restricted range of students' metacognitive ability. Following is a discussion for each grouping on what changes might strengthen future studies of this type.

Tutorial Concerns

Watson noted three concerns with the tutorial in the original lesson. The first being that the posttests are divided—one is about halfway through the lesson and another is at the end. He stated that some of the original subjects were confused because they did not anticipate two sets of questions. In the replication study, no students voiced confusion with the two posttests, although the researcher would not know if there were unvoiced concerns.

Watson also observed that a single-page screen, designed as an introduction to the lesson, confused some students by posing two guiding questions that were later answered as the students proceeded through the tutorial (see Figure 2). In the replication study, several students also raised their hands and asked if they were to answer the questions. Watson points out that the students' metacognitive awareness levels might have been altered as a result of this experience.

Watson's third concern was whether the metacognitive probe screen might be considered a metacognitive prompt as well. The metacognitive probe asked the user to reflect

on their knowledge, and self-assess what their posttest score might be. It was unknown what affect this screen had on the metacognitive-related statistical results in the original and replication studies. In addition, the study added another, more refined probe, right before the second posttest. The second probe may have had additional affects on students' metacognitive awareness levels.



Figure 2. Guiding question tutorial screen.

Posttests Concerns

There are two posttest concerns. The first deals with the reliability of the posttests and the second with the apparent differences in equality level of the posttests. Even with the addition of eight posttest questions, the reliability of the posttest was .62. While not as high as the .67 seen in the original study, it is interesting to note that if just the original posttest questions had been used this time, the reliability of the posttest would have been .50.

As mentioned previously, students did moderately well at predicting how they would do on the first posttest, but not the second one. The researcher realized that there might be a difference between the two posttest and ran a paired-samples *t* test. The mean on the first posttest was .72 and the mean on the second posttest was .50. This difference is significant at the .000 level. Clearly the students performed better on the first posttest. Although this difference did not show up in the original study, it is not surprising to the researcher. The first posttest covers material about states of matter—a topic that can be seen to have black

and white boundaries and the second posttest covers types of clouds. While fairly easy to understand the difference between a solid and a liquid, it is much harder for many people to distinguish types of clouds formations.

When the researcher realized that there was a significant difference between the two posttests, all previous data were analyzed again. This time the mean of the individual posttests was used as the dependent variable. The only significant difference that appeared was a gender difference on the $2 \times 2 \times 2$ between-subjects factorial ANOVA from question three. Boys mean score on the first posttest was 6.02 and girls mean score was 5.47.

Subjects' Metacognitive Abilities

Faithful replication of Watson's original study was both a strength and a limitation. In trying to find schools similar to Casillas, in family SES level and ethnicity, it is possible that many of the students who participated in the study were of a high metacognitive level already for their age group, since most students at the three schools come from high-SES families. Although there are more English Language Learners (ELL) at Parkview Elementary, the researcher did not need to drop any students solely because of ELL status. She believes that because the parent permission slip was not translated into Spanish, many of the ELL students at Parkview did not participate in the study because parents did not return consent forms. If the students were from a restricted range of metacognitive ability, then it is possible that they were too mature metacognitively to benefit from the low level prompt used in the study.

FUTURE RESEARCH

Although this study failed to replicate Watson's original findings, it did not disprove them, but instead illuminated some limitations to be corrected in future studies, namely a redesign of the tutorial, including the posttest, and a broader variation of subjects. It also raised a new question that is of particular interest to the researcher. How does a researcher calibrate a student's level of metacognitive ability and then match that ability to the appropriate type of prompt or intervention? Because of the role metacognition plays in problem solving and learning, and because computers are becoming more ubiquitous in homes and schools, understanding how metacognition can be taught or scaffolded with a computer is still a valid research concern.

Tutorial Redesign and Subject Selection

Beginning with the end in mind, using a highly reliable and well-validated posttest would be the first step to redesigning the tutorial. It is likely that this would also eliminate the dichotomous nature of the content used in the original and replication studies. Once the assessment piece is determined, the tutorial content should be designed to teach students the material covered on the posttest. Since it is possible that the act of taking the first posttest activated some students' metacognitively, the researcher suggests only one posttest at the end of the lesson to help eliminate possible variance.

In addition, future research should ensure that there is a sufficient range of metacognitively aware students in each group. There are two ways to ensure this range. First, students can be recruited from varying grade levels. Fifth graders (10-12 years old) are right on the cusp of what Piaget called the Phase of Formal Operations (Boyle, p. 91). Using students who are slightly younger and older (4th grade – 7th grade) would probably give a better picture of what is going on. Another way to vary the range of student's metacognitive levels is to select students from varying SES backgrounds.

Isolated Redesign

It is also very possible that the original and replication studies were too broad in scope. While both studies were well-planned, true experimental designs, each attempted to answer several research questions. Future researchers should consider isolating aspects such as gender and motivation.

Metacognitive Level and Appropriate Interventions

In 1985 Allen and Merrill suggested that one way to support achievement for students who are unable to determine which learning strategy to use is through the use of computer-based instruction systems that guide "the learner in the optimal use of previously-acquired internal processing strategies" (p. 5). Yet, researchers have only recently begun to study how to calibrate a student's metacognitive level and then correctly assign the appropriate intervention. While the IMSR can show a student's preference towards a metacognitive style, it is not currently used to prescribe "metacognitive treatments." Howard and colleagues write, if the IMSR's metacognitive styles can be further delineated and understood, "it might

be possible to train students to habitually use particular regulatory behaviors” (Howard, et. al, 2001a, n.p.).

This replication study attempted to add to the growing body of research that deals with understanding student’s metacognitive abilities and ways that metacognition can be fostered using computers. Although Watson’s original findings were not supported, the questions Watson raised are still valid and worth studying. Researchers today are trying to determine how to calibrate a student’s level of metacognition, and then assign appropriate tasks to foster metacognitive development. It is important to remember that Watson’s study was the only one the researcher could find that raised the question about how metacognitive prompts increase metacognitive awareness. In addition, Watson’s was the only study that raised the question about possible metacognitive differences between the genders. Both these issues are worth remembering as researchers continue to study metacognition in students.

REFERENCES

- Allen, B. S., & Merrill, M. D. (1985). System-assigned strategies and CBI. *Journal of Educational Computing Research*, 1(1), 3-21.
- Boyle, D.G. (1969). *A student's guide to Piaget*. Oxford: Pergamon Press.
- Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology: educational design and cognitive science* (pp. 77-165). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, L., & Brown, A. (1984). Metacognitive skills and reading. In P. D. Pearson, M. Kamil, R. Barr, & P. Mosenthal (Eds.) *Handbook of reading research* (pp. 353-394). New York: Longman.
- Caforti, N. (1996). *Examination of computer software in relation to gender differences*. Retrieved August 28, 2002, from Northeastern Illinois University, Department of Computer Science Web site: <http://www.math.luc.edu/ethics96/papers/caforti.txt>
- Chula Vista Elementary School District. (n.d.) *Our district*. Retrieved April 16, 2004, from <http://www.cvesd.k12.ca.us/cvesd/superin/ourdistrict.html>
- Clark, C. (2004, February). The boomburbs. *San Diego Magazine Online*. Retrieved April 16, 2004, from: <http://www.sandiegomag.com/issues/february04/featurea0204.asp>
- CTB McGraw-Hill. *California English Language Development Test*. Retrieved April 16, 2004 from <http://www2.ctb.com/state/CA/celdt/index.shtml>
- Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalfe & A. P. Shimamura (Eds.) *Metacognition: Knowing about knowing* (pp. 207-226). Cambridge, MA: MIT Press.
- Davidson-Shivers, G., Shorter, L., Jordan, K., & Rasmussen, K.L. (1999). Learning strategies and navigation decisions of children using a hypermedia lesson. *Journal of Educational Multimedia and Hypermedia* 8(2), 175-188.
- Delclos, V. R., & Harrington, C. (1991). Effects of strategy monitoring and proactive instruction on children's problem-solving performance. *Journal of Educational Psychology*, 83, 35-42.
- Desoete, A., Roeyers, H., & Buyse, A. (2001). Metacognition and mathematical problem solving in grade 3. *Journal of Learning Disabilities*, 34(5), 435-449.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). New York: Lawrence Erlbaum Associates.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911.

- Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. [Electronic version]. *Annual Review of Psychology*, 50, 21-45.
- Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development (4th ed.)*. Upper Saddle River, New Jersey: Prentice Hall.
- Harcourt Assessment, Inc. (2004) *California STAR program*. Retrieved on April 16, 2004, from <http://star.cde.ca.gov/star2002/help/AboutSTAR.html>
- Hennessey, M. G. (1999). *Probing the dimensions of metacognition: implications for conceptual change teaching-learning*. [Electronic version]. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching, Boston.
- Howard, B. C., McGee, S., Shia, R. M., & Hong, N. S. (2000, April). *Metacognitive self-regulation and problem-solving: Expanding the theory base through factor analysis*. Poster session presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. Retrieved online at <http://www.cet.edu/research/papers/CPhoward300.html>
- Howard, B. C., McGee, S., Shia, R. M., & Hong, N. S. (2001a). *Computer-based science inquiry: How components of metacognitive self-regulation affect problem-solving*. Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA. Retrieved online at <http://www.cet.edu/research/papers/CPhoward01.html>
- Howard, B. C., McGee, S., Shia, R. M., & Hong, N. S. (2001b). *The influence of metacognitive self-regulation and ability levels on problem solving*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA. Retrieved online at <http://www.cet.edu/research/papers/CPhoward201.html>
- Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, 22, 255-278.
- Kuhn, D. (2000). Metacognitive development. *American Psychological Society*, 9(5), 178-181.
- Kreutzer, M.A., Leonard, C., & Flavell, J.H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development*, 40 (1, Serial no. 159).
- Kurtz, B. E., & Borkowski, J. G. (1985, April 25-28). *Metacognition and the development of strategic skills in impulsive and reflective children*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Toronto, Canada.
- Kurtz, B. E., & Borkowski, J. G. (1987). Development of strategic skills in impulsive and reflective children: A longitudinal study of metacognition. *Journal of Experimental Child Psychology*, 43, 129-148.
- McKean, M., Allen, B., & Hoffman, B. (2000, April). *Sequential data analysis: Implications for assessment of usability in virtual museums*. Paper presented at

- the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Metcalf, J., & Shimamura, A. P. (Eds.). (1994). *Metacognition: Knowing about knowing*. Cambridge: MIT Press.
- Mevarech, Z. R. (1999). Effects of metacognitive training embedded in cooperative settings on mathematical problem solving. [Electronic version]. *The Journal of Educational Research*, 92(4), 195-205.
- Moran, C. (2004, February 7). A city divided—by Interstate 805. *San Diego Union-Tribune*. Retrieved April 16, 2004, from http://www.signonsandiego.com/news/education/chulavista/20040207-9999-_mz6m7chulavi.html
- Nelson, T. O., & Narens, L. (1994) Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.) *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.
- Packer, A. (2002). *Improving labor markets*. Retrieved April 15, 2004 from <http://www.scans.jhu.edu/NS/HTML/Index.htm>.
- Palinscar, A. M., & Brown, D. A. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117-175.
- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory Into Practice*, 41(2), 116-125.
- Roberts, D. F., Foehr, U. G., Rideout, V. J., & Brodie, M. (1999, November). *Kids & media @ the new millennium*. Menlo Park, CA: Kaiser Family Foundation.
- U. S. Department of Commerce. (2002). *A nation online: How Americans are expanding their use of the Internet*. Washington, DC: Author. Posted at: http://www.ntia.doc.gov/ntiahome/dn/nationonline_020502.htm.
- SCANS 2000 Center (2001). *SCANS skills-simple list*. Baltimore, MD: John Hopkins University, Institute for Policy Studies, SCANS 2000 Center. Retrieved April 15, 2004, from <http://www.scans.jhu.edu/NS/HTML/SkillsSimple.htm>.
- Schneider, W., & Lockl, K. (2002). The development of metacognitive knowledge. In T.J. Perfect & B.L. Schwartz (Eds.), *Applied Metacognition* (pp. 224-257). Cambridge, UK: Cambridge University Press.
- Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Schoenfield, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfield (Ed.), *Cognitive science and mathematics education* (pp. 189-215). Hillsdale, NJ: Lawrence Erlbaum Associates.

- School Accountability Report Card School Year 2001—2002* (Casillas Elementary, Hilltop Drive Elementary and Parkview Elementary) [SARC]. Chula Vista, CA: Chula Vista Elementary School District. Retrieved from http://www.cvesd.k12.ca.us/cvesd/inst_services/public/arc/arc.html
- Scruggs, T., Mastropieri, M.A., Monson, J., & Jorgenson, C. (1985). Maximizing what gifted students can learn: Recent findings of learning strategy research. *Gifted Child Quarterly*, 29: 181-185.
- Simons, P. R. J. (1996). Metacognition. In E. Corte & F. E. Weinert (Eds.), *International encyclopedia of developmental and instructional psychology* (pp. 436-444). Oxford, UK: Elsevier Science.
- Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, 27, 51-79.
- Sperling, R. A., Walls, R. T., & Hill, L. A. (2000). Early relationships among self-regulatory constructs: Theory of mind and preschool children's problem solving. *Child Study Journal*, 30(4), 233-254.
- Sundre, D. L. (2000). *Motivation scale background and scoring guide*. Retrieved from http://www.jmu.edu/assessment/wm_library_motscoreguide.pdf.
- Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, 82(2), 306-314.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Washington Roundtable. (1998, September). *What every high school graduate should know: A guide to a world of opportunity for people who have skills and knowledge*. Retrieved from <http://www.waroundtable.com/publications/publications.asp>
- Watson, J. B. (2001). The effect of embedded metacognitive cues and probes on use of learner control features in an on-line lesson for elementary students. *Dissertation Abstracts International*, 62 (07), 2394. (UMI No. 3020876)
- White, B. Y., & Frederickson, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16, 3-118.

APPENDIX A

PARENT PERMISSION SLIP

APPENDIX A

PARENT PERMISSION SLIP

University of San Diego/San Diego State University

Parental Permission/Informed Consent to Participate in Research

(The Effect of Embedded Metacognitive Prompts and Probes on Students' Awareness in an
On-line Lesson for Elementary Students)

You are being asked to allow your child to participate in a research study. Before you give your permission for your child to participate, it is important that you read the following information and ask as many questions as necessary to be sure you understand what your child will be asked to do.

Investigators:

Wendy Parcel, doctoral candidate. College of Education, Department of Educational Technology, San Diego State University and University of San Diego

Dr. Marcie Bober, supervisor. College of Education, Department of Educational Technology, San Diego State University

Purpose of the Study:

I am studying ways to improve learning in children through use of the computer.

Approximately 140 fifth grade students will participate in the study. They will be chosen randomly, which means by chance, to participate. However, they must speak English well because the computer program is available in English only.

Description of the Study:

If you agree to allow your child to participate, he/she will be asked to complete a short computer-based lesson about clouds and states of water that will take about 30-40 minutes during class time. Your child may also be asked questions about how he or she likes the computer lesson. Your child's participation in this study is completely voluntary and will not affect his/her grades in any way. Your child's school records will be accessed to obtain standardized test scores for the purposes of this study. The study will be conducted in May-June, 2004 within the Chula Vista Elementary School District.

Risks or Discomforts:

Your child may become tired or frustrated when trying to complete a study task or activity. If your child begins to feel uncomfortable he/she may stop participating in the study, either temporarily or permanently.

Benefits of the Study:

Your child will be given the opportunity to learn about clouds and states of water, as well as use a new computer program. I cannot guarantee, however, that you or your child will receive any benefits from this study.

The results of the study will, hopefully, aid those creating educational computer programs to provide the best possible environment for learning.

Confidentiality:

To protect your child's confidentiality, your child's name will not be shared with anyone, unless required by law. We will store information about how your child uses the computer program and may interview a small number of participants, approximately 30 students. Only the researcher will have access to this information.

Incentives to Participate:

Your child will not be paid for his/her participation, but may be offered a small incentive (such as a pencil) to try and perform well on the quiz.

Voluntary Nature of Participation:

Participation in this study is voluntary. Your decision of whether or not to allow your child to participate will not prejudice your future relations with San Diego State University, the University of San Diego or your child's present school. If you decide to allow your child to participate, you are free to withdraw your consent and to discontinue his/her participation at any time without penalty or loss of benefits to which you are otherwise entitled.

Questions about the Study:

If you have questions about the research now, please ask. If you have questions later about the research, you may contact Wendy Parcel at 619-585-6040.

If you have questions regarding your child's rights as a human subject and participant in this study, you may contact the Institutional Review Board at San Diego State University for information (619-594-6622 or irb@mail.sdsu.edu) and/or the Office of the Vice President

and Provost, University of San Diego 5998 Alcala Park, San Diego, CA 92110 (telephone: 619-260-4553).

Agreement:

The Institutional Review Boards at San Diego State University and the University of San Diego have approved this consent form.

Your signature below indicates that you have read the information in this document and have had a chance to ask any questions you have about the study. Your signature also indicates that you agree to allow your child to be in the study and have been told that you can change your mind and withdraw your consent to participate at any time. You have been given a copy of this agreement. You have been told that by signing this consent document you are not giving up any of your legal rights.

Name of Participant (please print)

Signature of Parent or Guardian

Date

Signature of Investigator

Date

Protocol #04-03-147X

APPENDIX B

STUDENT PERMISSION SLIP

APPENDIX B

STUDENT PERMISSION SLIP

San Diego State University

Assent to Participate in Research

(The Effect of Embedded Metacognitive Prompts and Probes on Students' Awareness in an
On-line Lesson for Elementary Students)

1. My name is Wendy Parcel.
2. We are asking you to take part in a research study. We are trying to learn more about how to make learning on computers easier and more fun for you.
3. If you agree to be in this study, you will be asked to spend about fifteen minutes learning about clouds and states of water on the computer.
4. You may get tired or frustrated during the computer lesson. If you do, please notify me right away, and you will be able stop.
5. By participating in this study, we hope you'll learn a little about clouds and states of water, as well as about how to use this computer software.
6. Please talk to your parents about this study before you decide whether to participate. We will also ask your parents if it is all right with them for you to take part in this study. If your parents say that you can be in the study, you can still decide not to participate.
7. You can ask me any questions that you have about this study and I will try to answer them for you. If you have questions that you think of later, you can call me at 619-585-6040.
8. Taking part in this study is up to you. No one will be upset if you don't want to participate. If you decide to participate, you can also change your mind and stop any time you want.

Please mark one of the choices below to tell us what you want to do:

_____ No, I do not want to be in this project _____ Yes, I want to be in this project.

Write your name here

Date

Project Representative

Date

APPENDIX C

IMSR

APPENDIX C

IMSR

How do You Solve Problems?

Please read the following sentences and circle the answer that best describes the way you are when you are trying to solve a problem. Think about a problem that you might see in a science or math class.

1. Think about when you have to solve a hard problem. What do you do before you start?
2. What do you do while you work the problem?
3. What do you do after you finish working the problem?

There are no right answers--please describe yourself as you are, not how you want to be or think you ought to be. Your teacher will not grade this.

1= never 2= seldom/ rarely 3=sometimes 4= often/ frequently 5= always

- | | | | | | |
|---|---|---|---|---|---|
| 1. I try to understand what the problem is asking me. | 1 | 2 | 3 | 4 | 5 |
| 2. I think of several ways to solve a problem and then choose the best one. | 1 | 2 | 3 | 4 | 5 |
| 3. I look back at the problem to see if my answer makes sense. | 1 | 2 | 3 | 4 | 5 |
| 4. I use different ways to memorize things. | 1 | 2 | 3 | 4 | 5 |
| 5. I think to myself, do I understand what the problem is asking me? | 1 | 2 | 3 | 4 | 5 |
| 6. I read the problem more than once. | 1 | 2 | 3 | 4 | 5 |
| 7. I think about what information I need to solve this problem. | 1 | 2 | 3 | 4 | 5 |
| 8. I use different learning strategies depending on the problem. | 1 | 2 | 3 | 4 | 5 |
| 9. I look back to see if I did the correct procedures. | 1 | 2 | 3 | 4 | 5 |
| 10. I think about how well I am learning when I work a difficult problem. | 1 | 2 | 3 | 4 | 5 |
| 11. I use different ways of learning depending on the problem. | 1 | 2 | 3 | 4 | 5 |
| 12. I go back and check my work. | 1 | 2 | 3 | 4 | 5 |
| 13. I read the problem over and over until I understand it. | 1 | 2 | 3 | 4 | 5 |
| 14. For this question, please circle number 2. | 1 | 2 | 3 | 4 | 5 |
| 15. I check to see if my calculations are correct. | 1 | 2 | 3 | 4 | 5 |
| 16. When it comes to learning, I can make myself learn when I need to. | 1 | 2 | 3 | 4 | 5 |
| 17. I ask myself how well I am doing while I am learning something new. | 1 | 2 | 3 | 4 | 5 |
| 18. I check my work all the way through the problem. | 1 | 2 | 3 | 4 | 5 |
| 19. I identify all the important parts of the problem. | 1 | 2 | 3 | 4 | 5 |

20. I try to understand the problem so I know what to do.	1	2	3	4	5
21. I think about all the steps as I work the problem.	1	2	3	4	5
22. I can make myself memorize something.	1	2	3	4	5
23. When it comes to learning, I know my strengths and weaknesses.	1	2	3	4	5
24. I pick out the steps I need to do this problem.	1	2	3	4	5
25. When I am done with my schoolwork, I ask myself if I learned what I wanted to learn.	1	2	3	4	5
26. I double-check to make sure I did it right.	1	2	3	4	5
27. For this question, please circle number 1.	1	2	3	4	5
28. I try to break down the problem to just the necessary information.	1	2	3	4	5
29. I use learning strategies without thinking.	1	2	3	4	5
30. When it comes to learning, I know how I learn best.	1	2	3	4	5
31. I ask myself if there are certain goals I want to accomplish.	1	2	3	4	5
32. I try more than one way to learn something.	1	2	3	4	5

The Inventory of Metacognitive Self-Regulation

Copyright © 1998 by Wheeling Jesuit University/ Center for Educational Technologies. All rights reserved.

APPENDIX D

PERMISSION TO USE THE IMSR

APPENDIX D

PERMISSION TO USE THE IMSR

Hi Wendy,

I'd be happy to share it with you. All I ask is that you send me a copy of your results when you are done. BTW: If you're working with folks over 18, you may want to use the Metacognitive Awareness Inventory instead:

See Schraw , G., & Dennison (1994). Assessing metacognitive awareness.
Contemporary Educational Psychology, 19, 460-475.

You can contact one author: Rayne Sperling at rsd7@psu.edu. She's a friend. Tell her I sent you her way.

Best,
Bruce Howard, Ed.D.
Curriculum Director
Challenger Learning Center
Wheeling, WV 26003

-----Original Message-----

From: Wendy J. Parcel [<mailto:wendy.parcel@suhsd.k12.ca.us>]
Sent: Monday, January 27, 2003 5:42 PM
To: howard@cet.edu; mcgee@cet.edu
Subject: Inventory of Metacognitive Self-Regulation

I am a doctoral student in Educational Technology at the University of San Diego/ San Diego State. I recently read your paper that was presented at AERA in 2000. I am interested in using your Inventory of Metacognitive Self-Regulation when conducting my dissertation study. I'm going to be looking at user control in a hypermedia learning environment with metacognitive prompts for the experimental group. I'd like to give the students a test of metacognition after the treatment.

I'm just beginning this process and I am sorry if people normally don't write to the author and I've just not discovered the protocol yet. I've looked at Burros, but don't see the test. Is there a way I can order the test.

Thank you for taking the time to read this and help.

Wendy Parcel
Sweetwater Union High School District

APPENDIX E

STUDENT OPINION SCALE

APPENDIX E

STUDENT OPINION SCALE

Please think about both of the quizzes that you just completed. Mark the answer that best represents how you feel about each of the statements below.

1. Doing well on the quizzes was important to me.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

2. I engaged in good effort throughout the quizzes.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

3. I am not curious about how I did on the quizzes relative to others.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

4. I am not concerned about the score I receive on the quizzes.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

5. These were important quizzes to me.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

6. I gave my best effort on the quizzes.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

7. While taking the quizzes, I could have worked harder.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

8. I would like to know how well I did on the quizzes.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

9. I did not give the quizzes my full attention.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

10. While taking the quizzes, I was able to try hard until the end of each quiz.

Strongly Disagree | Disagree | Neutral | Agree | Strongly Agree

APPENDIX F

INTERVIEW QUESTIONS

APPENDIX F

INTERVIEW QUESTIONS

Direction to Interviewer: Make sure the student is comfortable

Say, "We won't share what you tell us with anyone, not even your parent or teacher. Anything you say to us is private--it is just between you and me."

Questions

Say to the student. "When you began the tutorial you learned a little bit about how water can be a solid or a liquid. Then there was a short test on the material. Do you remember?

Remember when you saw one of these screen the first time? (Hold up a print out of the prompt/no prompt screen.) It was right before the first test.

When you saw this screen, did you click on the back button?

o Yes

o No

Why

Can you think of anything else?

Once again show the student the picture of the prompt/no prompt screen.

What did you think about when you saw this screen?

Can you think of anything else?

APPENDIX G

INSTRUCTION FOR SUBJECTS

APPENDIX G

INSTRUCTION FOR SUBJECTS

Hello.

Thank you for helping us out by trying this new lesson software. The lesson is about water, weather and clouds. The software is easy to use, and there are a few questions to test to see what you have learned. Just do the best you can to learn about water, weather and clouds. After you watch this video, someone will help you log into the program and get started. We are here to answer any of your questions and to make sure everything runs smoothly.

You should see a form on the desk in front of you. Please read it and if you agree with the form, sign the bottom. We just want to make sure you're ready to try this new software.

Again, if you have any questions, stop and raise your hand.

When you are ready, please let someone know.

APPENDIX H

STUDENT LOGIN LETTERS

APPENDIX H

STUDENT LOGIN LETTERS

97

5

This is a very important scientific experiment. Please help me by not telling anyone about what you read on this paper, or about the computer tutorial.

Remember--don't tell anyone.

98

3

This is a very important scientific experiment. Please help me by not telling anyone about what you read on this paper, or about the computer tutorial.

For every question you get right on the quizzes, you will get a nickel. Wendy will return next week to let you know how you did on the quizzes.

Remember--don't tell anyone.